

No.	Score	Match	Length	DB	ID	Description
1	124.4	5.8	126	4	US-09-172-841-2	Sequence 2, Appl1
2	82.2	3	2237	3	US-08-914-999-7	Sequence 7, Appl1
3	67.2	3.1	2481	4	US-08-899-578-1	Sequence 1, Appl1
4	50.8	2.4	2085	2	US-08-283-917-8	Sequence 8, Appl1
5	50.8	2.4	2085	2	US-08-961-716-8	Sequence 8, Appl1
6	50	2.3	7218	1	US-08-232-463-14	Sequence 14, Appl1
7	47.8	2.2	2152	1	US-08-188-582-17	Sequence 17, Appl1
8	47.8	2.2	2152	1	US-08-664-715-17	Sequence 17, Appl1
9	39.8	1.9	7218	1	US-08-232-463-14	Sequence 14, Appl1
10	36.6	1.8	433	1	US-08-664-596B-11	Sequence 1, Appl1
11	38.6	1.8	433	1	US-08-738-367-1	Sequence 1, Appl1
12	38.6	1.8	733	1	US-08-738-367-6	Sequence 6, Appl1
13	36.6	1.8	3380	2	US-09-156-425-1	Sequence 1, Appl1
14	36.6	1.7	18596	4	US-09-318-448-11	Sequence 11, Appl1
15	36.2	1.7	7886	2	US-08-751-189-2	Sequence 2, Appl1
16	36.2	1.7	7886	2	US-09-060-836-2	Sequence 2, Appl1
17	36.2	1.7	7886	4	US-09-184-445-2	Sequence 2, Appl1
18	35.8	1.7	3434	4	US-09-439-313-776	Sequence 476, Appl1
19	35.8	1.7	5829	4	US-09-439-313-773	Sequence 473, Appl1
20	35.4	1.6	7881	2	US-08-751-189-1	Sequence 1, Appl1
21	35.4	1.6	7881	2	US-09-060-836-1	Sequence 1, Appl1
22	35.4	1.6	7881	4	US-09-184-445-1	Sequence 1, Appl1
23	34.8	1.6	3465	4	US-08-914-999-5	Sequence 5, Appl1
24	34.6	1.6	1848	1	US-08-313-553-10	Sequence 10, Appl1
25	34.6	1.6	1848	3	US-08-762-993-10	Sequence 10, Appl1
26	34.6	1.6	2246	4	US-09-032-742-31	Sequence 21, Appl1
27	34.6	1.6	2246	4	US-09-032-742-21	Sequence 21, Appl1

28	34.6	1.6	2246	4	US-09-032-742-22	Sequence 22, Appl
29	34.6	1.6	2245	4	US-09-032-742-24	Sequence 22, Appl
30	34.6	1.6	2246	4	US-09-032-742-25	Sequence 22, Appl
31	34.4	1.6	649	4	US-09-385-982-336	Sequence 35, Appl
32	34.2	1.6	7042	4	US-09-092-508-1	Sequence 1, Appl
33	34.2	1.6	7042	4	US-09-435-115-1	Sequence 1, Appl
34	34.2	1.6	7075	4	US-09-092-508-15	Sequence 15, Appl
35	34.2	1.6	7075	4	US-09-435-115-15	Sequence 15, Appl
36	33	1.5	1758	3	US-09-191-171-6	Sequence 6, Appl
37	33	1.5	1758	4	US-09-385-982-1	Sequence 1, Appl
38	32.6	1.5	1368	4	US-09-035-382-1	Sequence 1, Appl
39	32.6	1.5	758	3	US-08-506-769-126	Sequence 126, Appl
40	32.2	1.5	758	3	US-08-506-616-126	Sequence 126, Appl
41	32.2	1.5	758	3	US-08-639-075A-126	Sequence 126, Appl
42	32.2	1.5	758	4	US-09-012-431-126	Sequence 126, Appl
43	32.2	1.5	758	4	US-09-012-692-126	Sequence 126, Appl
44	32.2	1.5	758	4	US-08-906-613-126	Sequence 126, Appl
45	31.8	1.5	567	3	US-08-923-454A-22	Sequence 22, Appl

ALIGNMENTS

```

US-09-172-841-2
SEQUENCE 1
: Sequence 2, Application US/09172841
: Patent No. 6232081
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeffrey W.
: APPLICANT: Elledge, Stephen J.
: TITLE OF INVENTION: F-BOX PROTEINS AND GENES
: FILE REFERENCE: BCM-0310
: CURRENT APPLICATION NUMBER: US/09/172,841
: CURRENT FILING DATE: 1998-10-15
: EARLIER APPLICATION NUMBER: 08/951,621
: EARLIER FILING DATE: 1997-10-16
: NUMBER OF SEQ ID NOS: 60
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 126
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-172-841-2

```

```

Query Match 5.8%; Score 124.4; DB 4; Length 126;
Best Local Similarity 99.2%; Pred. No. 4.8e-31;
Matches 125; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Oy 511 ctgcagcctcgaggatgtgatacatatcgcttggaacattctctcatacctgatacca 570
|||||
Db 1 ctgcagcctcgaggatgtgatacatatcgcttggaacattctctcatacctgatacca 60
|||||

Oy 571 tcaatattgtctctggaactgtgtgcaagaaatgttaccgaagtgaacctgtatgcatg 630
|||||
Db 61 tcaattgtctgtgaactgtgtgcaagaaatgttaccgaagtgaacctgtatgcatg 120
|||||

Oy 631 ctgttgg 636
|||||
Db 121 ctgtgg 126

RESULT 2
US-08-914-999-7
; Sequence 7, Application US/08914999
; Patent No. 6346406
; GENERAL INFORMATION:
; APPLICANT: Ryazanov, Alexey G.
; APPLICANT: Halt, William N.
; APPLICANT: Pavur, Karen S.
; TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
; TITLE OF INVENTION: AND METHODS OF USE THEREFOR
; NUMBER OF SEQUENCES: 25
;

```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2237 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ORGANISM: Dictyostelium discoideum
US-08-914-999-7

```

```

Query Match          3.8%; Score 82.2; DB 4; Length 2237;
Best Local Similarity 47.4%; Pred. No. 2.3e-16;
Matches 286; Conservative 0; Mismatches 308; Indels 9; Gaps 1;

QY 959 tggatgcaagcagatctccacagccacaggtcagtcctcctcgtcctcagatgagtg 1018
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1490 TCGATGTTGTTCAACCTTGAAGGTCATGAAGGTCAGGTCGATCAATTTGTTATATG 1549
QY 1019 agagagtcatacaacagatcagatccacggtcagatgagtgagtgatgaaatacag 1078
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1550 ATCAATATTTTGTAGTGTGTCATCATCAATTAAGTTTGGATTTAAAGAAAT 1609
QY 1079 gtgaatgtcaaacgcgtgtatccacctgtgaaagcagtcctcagctcgttccaata 1138
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1610 TAAGATGATTTTGTAGTGTGAGGTCATGATTAACCTGTCATACAGGTTCTATTTGATG 1669
QY 1139 atgcatgagtggtgacccgtcccaagaatcgctcagtcgtatgagatagccccc 1198
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1670 ATAATATTTTGTAGTGTGTCCTCTGCAAACTATCAAAAGTTTGGGATTTGAAAACTT 1729
QY 1199 caactgacatacccccgcgaggggtgctgctgagacacgcagctgctgcaatgtgtag 1258
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1730 TGAATGTAAATATATACCTTGAAGTCATGCCAGA-----GCCGTCAAAACACTTT 1780
QY 1259 acttgatgacaagtaactgttctcgtcattcgtgggagatagacataaagtgtagaca 1318
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1781 GTATATCTGTCATATTTATTTAGTGTTCATAAATGATTAAGTCAAGGTTTGGGATT 1840
QY 1319 caagtaactgtgaaattgtgaaagccttaaaatgagacaacagagcagctgctgttcg 1378
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1841 TGAAGACTTTTCGTTGTAACACACCTCAAAAGGTCATACCAATGGGTCACCACTATCT 1900
QY 1379 agtcaagagcagcgtgtagtgagtgctcattgacaacacatacagaatgtaggaca 1438
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1901 GTATATTAAGTACCAATCTACAGTGTGCTCTATGATTAAGACTATTAAGGTTTGGAAAT 1960

```

```

QY 1439 tagaatggtgcatggtttacagagtgtagagggccatgggaattggtcgttatc 1498
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1961 TAAAGAGTTTGAATGTTCCCTTACTTAAAGGCGCATGATGAGTGGTGAACATATGG 2020
QY 1499 gattgataacaagagatagtcagtgaggccctatgatagaataaataaagtgtagac 1558
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 2021 TAATTGTGATTAATTTATTTACTCTGCTAGTACGATATATACATTAATAATTTGGGATT 2080
QY 1559 ttg 1561
DB 2081 TAG 2083

```

```

RESULT 3
US-08-899-578-1
Sequence 1, Application US/08899578
Patent No. 6087153
GENERAL INFORMATION:
APPLICANT: Greenwald, Iva
APPLICANT: Hubbard, E. Jane
TITLE OF INVENTION: SEL-10 AND USDS THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,578
FILING DATE: 24-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/53200/JFW/AKC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 278-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2481 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 91..1854
US-08-899-578-1

```

```

Query Match          3.1%; Score 67.2; DB 3; Length 2481;
Best Local Similarity 50.6%; Pred. No. 2.1e-11;
Matches 162; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 1273 tacatgttctgcatactcgtggagatacaagatgataagacacagtaactgtgaa 1332
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1018 TATATTTAGGGGGGTCCTCATAGAACTGTAAAGTTTGGAGTACTGTAGATGTTCA 1077
QY 1333 ttgtgaagaccttaaatgagacaacagagcaltgctgtgttgtagtagagagacag 1392
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1078 CTTCTTCATACACTTCAAGACATCTTCCACTGTCATGATGATGATGATGATGATGATG 1137
QY 1393 ctgtagtagtggtcctcagacaacatcagatgtaggacatagatgtggtgca 1452
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 1138 ATACTTGTCTGATCAACGAGATACCACTTCTGATGAGAGTACCAATCCGACGT 1197

```


RESULT 6
US-08-232-463-14/c
:Sequence 14, Application US/08232463

[illegible]

DB 1625 ATTCTATTATGTTGCTACGGGCTCTGCAGACAGAACTGTGGGCTGTGGAGCTCCCTGA 1684
QY 1325 ctgtgtaatttgaagacctaaatggacacaaagagcatgtcctgtttgcagta-- 1382
DB 1685 ATGGTACTGTGTAAGGATTTCTTCTACGTGACACAAAGGCAATTCCTTCTGACATTTT 1744
QY 1383 ----caggacacagctgtgtagtgcgtcctgcacacatactcagattatggagca 1438
DB 1745 CTCCTCAATGGAGATTCCTCGCTACAGAGCAACAGATGGCAGATGCTTCTTTGGGATA 1804
QY 1439 tagaatgtgctcagtttaagagtgtaggaagccatgaggaattgtgcgtgtatc 1498
DB 1805 TTGGACATGCTTGTGATGTTGGAGATTAAAGGCCACACATGATACAGTCTGTTCACTTAA 1864
QY 1499 gattgata 1507
DB 1865 GGTTAGTA 1873

RESULT 9

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ. ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-fls
; US-08-232-463-14

Query Match 1.9%; Score 39.8; DB 1; Length 7218;
Best Local Similarity 3.08; Pired. No. 0.042;
Matches 14; Conservative 210; Mismatches 167; Indels 0; Gaps 0;

QY 1714 gatccagctcccaagctgaaccccccggtccctctcgaacatacacctacatctco 1773
DB 1052 GAGGAGCTTGCGATTT 1111
QY 1774 agataataaccac 1833
DB 1112 TT 1171
QY 1834 aacgtatcgcacataccagatgagcaacacacacacacacacacacacacacacacac 1893
DB 1172 TT 1231
QY 1894 cccctgagctccgag 1953
DB 1232 TT 1291
QY 1954 tcggcccaagagctcactcagacacacacacacacacacacacacacacacacacacac 2013
DB 1292 TT 1351
QY 2014 ctctatcaattgtaattggaactttaacccctcctcctcctcctcctcctcctcctc 2073
DB 1352 TT 1411
QY 2074 ctgcacctagtttttccatgtgtcaga 2104
DB 1412 TT

RESULT 10

US-08-664-596B-1
; Sequence 1, Application US/08664596B
; Patent No. 5807703
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Lavallee, Edward
; APPLICANT: Racine, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Evans, Cheryl
; APPLICANT: Bowman, Michael
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/664,596B
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ. ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 base pairs
; TYPE: nucleic acid

OY 1516 atagtcagtgggcccatagatgaaaaaatlaaaqgtgttgatccttctgcctgc 1568
|| || | | | | | | | | | | | | | |
Db 227 CTATATTCCGACATCGTGGACTCTACAGTAGAGGTGTGCCTGTGTTTCCTGC 279

```

RESULT 13
US-09-156-425-1
: Sequence 1, Application US/09156425B
: Patent No. 5962671
: GENERAL INFORMATION:
: APPLICANT: Baker, Brenda F.
: APPLICANT: Coswert, Lex M.
: TITLE OF INVENTION: ANTISENSE MODULATION OF FAN EXPRESSION
: FILE REFERENCE: RTS-0009
: CURRENT APPLICATION NUMBER: US/09/156,425B
: CURRENT FILING DATE: 1998-09-18
: NUMBER OF SEQ ID NOS: 47
: SEQ ID NO 1
: LENGTH: 3380
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (13)..(2766)
: US-09-156-425-1

```

Query Match	1.8%;	Score 38.6;	DB 2;	Length 3380;
Best Local Similarity	51.4%;	Pred. No. 0.061;		
Matches	89;	Conservative	0;	Mismatches 84;
			Indels	0;
			Caps	0;

[illegible]

```

RESULT 14 448
US-09-318-448-11
; Sequence 11, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Steenboos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318, 448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 18596
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-318-448-11

```

Query Match	1.7%;	Score 36.6;	DB 4;	Length 18596;
Best Local Similarity	56.1%;	Pred. NO. 0.89;		
Matches 69;	Conservative	0;	Mismatches 54;	Indels 0;
				Gaps 0;

Dy 750 tcccaactccttttatagagcaattatccctaanaattatacaagacattgagacaataga 809
||| ||| ||| | | | | |
Db 16784 tgccagcccttgatatagctatttctgtgaagacacaccctcgacttcgggttaaa 16843

Dy 810 atctaattgagatgttgaagacatagtttacagagaattcactgcggaagtgaacaag 869
||||| | ||| | ||| | | ||| | ||| |||||
Db 1684 atctaaattattaagaattaagtagagataaacgttcattgatattgtctaaaagaatcaag 16903

QY	870	caa	872
Db	16904	taa	16906

```

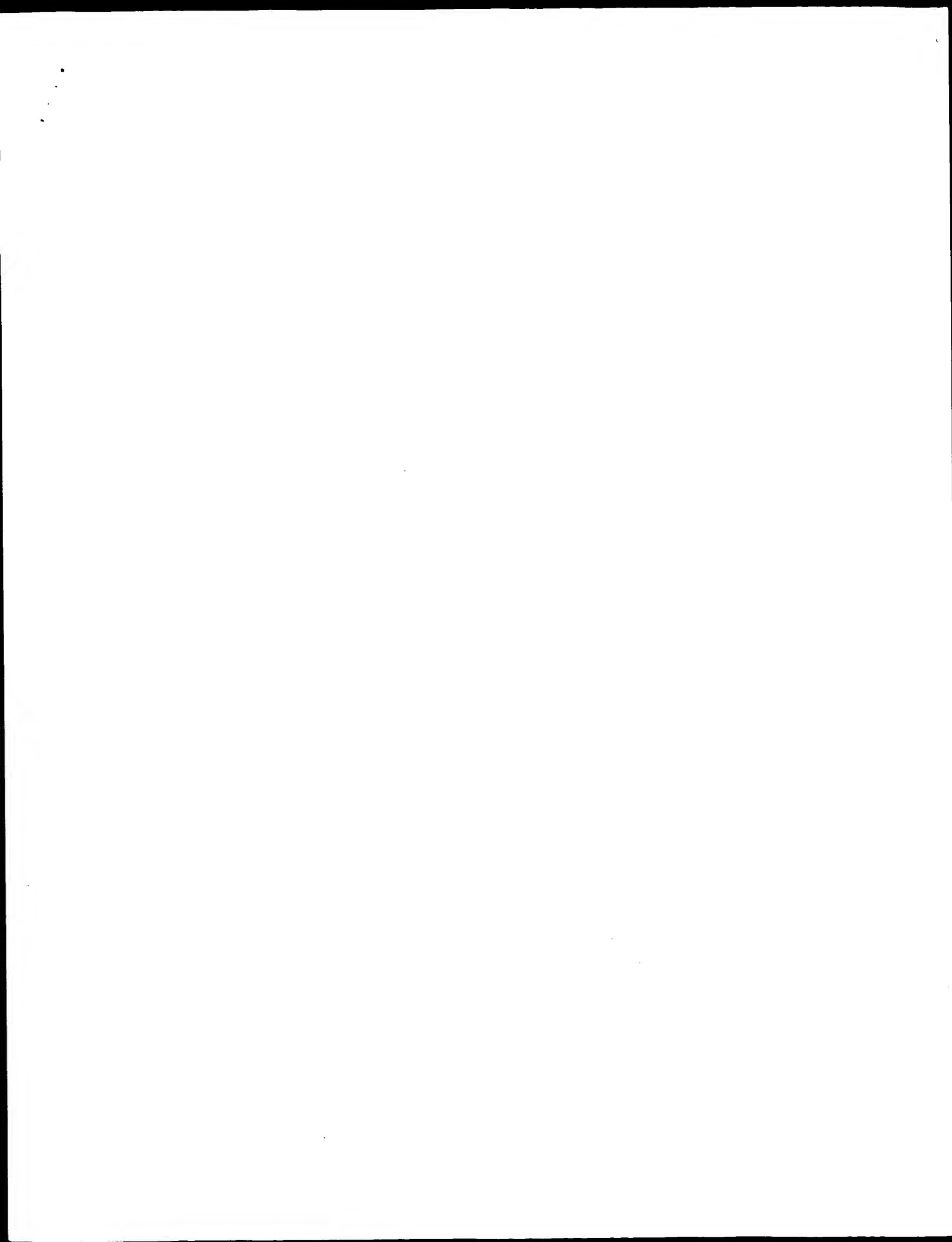
RESULT 15
US-08-751-189-2
: Sequence 2, Application US/08751189
: Patent No. 5919656
GENERAL INFORMATION:
APPLICANT: Harrington, Lea A.
APPLICANT: Robinson, Murray O.
TITLE OF INVENTION: No. 5919656el Genes Encoding Telomerase Protein
TITLE OF INVENTION: 1
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen, Inc.
STREET: 1840 De Havilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,189
FILING DATE: 15-NOV-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oleski, Nancy A.
REGISTRATION NUMBER: 34,688
REFERENCE/DOCKET NUMBER: A-433
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 7886 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-751-189-2

```

Query Match	1.7%;	Score 36.2;	DB 2;	Length 7886;
Best Local Similarity	47.6%;	Pred. No. 0.67;		
Matches 107;	Conservative	0;	Mismatches 118;	Indels 0;
			Gaps	0

QY	1365	caagcgccttttcacaaaggaagcgtgtagtagtgatcgtacatccgaacaacat	1422
Db	6353	CACTGGCTGTGGTGGACCAAGACACACTCCTGGTCTCTCTGAGTAGAGCTTGT	6412Z
QY	1425	cagattatggagacatgaatctgtgtcagtttaccagtggttaagaagccatgaagaat	1488
Db	6413	GGGACTCTGGAAACCCAGAGGAGGAGCACTTGGCCAGTTCTCAGGGCCACCAAGTGC	6472Z
QY	1485	ggtgcgtgtgattcgaattgataaagaagatagtcagttgggcctatagtgaanaat	1544
Db	6473	CGTGAGCCGCCGTTGGTGTCTGTGGAGAGACACTTGTATCTGTGAGACCCGAGATGAGCACTT	6532Z
QY	1545	taaaagtgtggaatctctgtgtgcgctttagaaccccgctgctctgc	1589
Db	6533	GAAAGTGTGGGACCATCAGAGGCTGTGGAGTCGACCAAGCATCCCTGC	6577

Search completed: May 11, 2002, 09:19:00
Job time: 13474 sec



Qy	1801	acttccccgagagcccatlaaagtctgcgtaattcaagctatctcgcaatccagagatgagc	1860
Db	1801	acttccccgagagcccatlaaagtctgcgtaattcaagctatctcgcaatccagagatgagc	1860
Qy	1861	aacaaacagtaacaatcaatcaactacgtgccagtttccctgtgactagccgagagcaggtctt	1920
Db	1861	aacaaacagtaacaatcaatcaactacgtgccagtttccctgtgactagccgagagcaggtctt	1920
Qy	1921	tgaagactccgtctgttggaacacagctgagctcgcagctcgccccagagacggtctatctagcaca	1980
Db	1921	tgaagactccgtctgttggaacacagctgagctcgcagctcgccccagagacggtctatctagcaca	1980
Qy	1981	actgacgtcttcagtgctgcgtcatcagaagaagtctcttcatcoaattgtgaatgttgaac	2040
Db	1981	actgacgtcttcagtgctgcgtcatcagaagaagtctcttcatcoaattgtgaatgttgaac	2040
Qy	2041	ttttaaacctccccctccctcctcctcttcaacctcgcagacctgatttctccattgtgtc	2100
Db	2041	ttttaaacctccccctccctcctcctccttcaacctcgcagacctgatttctccattgtgtc	2100
Qy	2101	cagacaagaagtgactataataatataatttagtgcttgcgcagaaaaaanaaa	2151
Db	2101	cagacaagaagtgactataataatataatttagtgcttgcgcagaaaaaanaaa	2151

RESULT: 2			
AX057166			
LOCUS	AX057166	2151 bp	DNA
DEFINITION	Sequence	42 from Patent WO0075184.	linear
ACCESSION	AX057166		
VERSION	AX057166.1	GI:12309979	

SOURCE	ORGANISM
human.	Homo sapiens
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
	Mammalia: Eutheria: Primates: Catarrhini: Homiinae: Homo.
REFERENCE	1 (bases 1 to 2151).
AUTHORS	Zhang, H., Tsvetkov, L. M. and Kondo, T.
TITLE	Modulation of protein levels using the scf complex
JOURNAL	Patent: WO 0075184-A 42 14-DEC-2000;

FEATURES	
source	location/Qualifiers
	1. .2151
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
BASE COUNT	628 a 467 c 513 g 543 t
ORIGIN	

Query Match	100.0%	Score 2151;	DB 6;	Length 2151;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2151; Conservative	0;	Mismatches	0;	Gaps 0;

QY	1	tcgcaatcctgagccctggcaccgaaggggagcccgcgagagcgagaccatctgccc	60
Db	1	tcgcaatcctgagccctggcaccgaaggggagcccgcgagagcgagaccatctgccc	60
QY	61	tcggcgatctatcggaccggccggcgagcggtgtctgcaagaagagcgacctaagtltatgcat	120
Db	61	tcggcgatctatcggaccggccggcgagcggtgtctgcaagaagagcgacctaagtltatgcat	120
QY	121	tctctcagaagagaagaactgtataataatgcygaaccccttagaagatataccagaagaag	180
Db	121	tctctcagaagagaagaactgtataataatgcygaaccccttagaagatataccagaagaag	180
QY	181	aattcactatagaagaaatatacaagcgtgtccagagactctgtcttaaaccaagaagaacgta	240
Db	181	aattcactatagaagaaatatacaagcgtgtccagagactctgtcttaaaccaagaagaacgta	240
QY	241	tgttttagcaagcactgtcatatgaagacttgaagaattgtgtgycgcaaaacaaacttgcacat	300
Db	241	tgttttagcaagcactgtcatatgaagacttgaagaattgtgtgycgcaaaacaaacttgcacat	300

QY	301	ggcacttcacagatgatgtgtgcccaagaacggaaacctcttcgaacgatagtaaaaggaa	360
Db	301	ggcacttccagatgatgtgtgcccaagaacggaaacctcttcgaacgatagtaaaaggaa	360
QY	361	aaggaactgtgtgtcaaaactcttbaagcagtggtcagaatcaga tcaagtggaaattgtg	420
Db	361	aaggaactgtgtgtcaaaactcttbaagcagtggtcagaatcaga tcaagtggaaattgtg	420
QY	421	gaacgtcttatatcccaaatgtgtcatctaaccaatgggcacataaactcgtatcttaa	480
Db	421	gaacactcttatatcccaaatgtgtcatctaaccaatgggcacataaactcgtatcttaa	480
QY	481	ccatgtttgcagaagaatctcataactgtctgcacgctcgaggatctggatcatatgcct	540
Db	481	cctatgtttgcagaagaatctcataactgtctgcacgctcgaggatctggatcatatgcct	540
QY	541	ggagaactcttgctacactggatagcgaatacactatagtgtctgtgaacttggttgaag	600
Db	541	ggagaactcttgctacactggatagcgaatacactatagtgtctgtgaacttggttgaag	600
QY	601	gaatgtacacgaatgaacctctgatatgacatgctgttggaaagaactctatcgagaagaatg	660
Db	601	gaatgtacacgaatgaacctctgatatgacatgctgttggaaagaactctatcgagaagaatg	660

87 001 aagacacagatcttcctgcgaggaagagcttcctccgcaaacagcaagaaaggaaggaggacagcattatctc 720
 Db 661 AGGACAGATTCTCTGTGAGAGAGCTCGACAAAGGAAGAGATGGGACAGATATTATTTC 720
 Qy 721 aaaaacaacactcctgcagcgggaatgcctctcccaactcttttatagacacttalcct 780
 Db 721 AAAAAACAACCTCCGAGCGGGGATGCTCTCCCAACTCTTTTATAGAGACCTTATTC 780

Accession	Sequence	Length
OY	aaatctacacagaacatttggagacaaatgaactctaatctggagatcttggagaaagacaaagtctt	840
Db	781 AAAATTATACAAAGACATTGGACACATATACATCTATTGGAGATGTGGACAGACATAGTTTA	840
OY	841 cagaatctcaactctgcggaaagtggaaacaaagagatttactctttacagtatgagat	900
Db	841 CAGAGATTCCTACTGCGGAAGTGAACCAAGCAAGAGAGTTTACTTTACAGTATGATGAT	900

Accession	Sequence	Length
901	CAGAAAATGTAAGGGGCGCTTCGAGACACACAAATCAAGATCTGGGATATAAAACACATTG	960
961	gaatgcaagcgaaattctctacagcgccatacagaatttcagttctctgltccagatgatgag	1020
961	GAATGCAAGCGAATTCTCTACAGCGCCATACAGAGTTCACTGCTCTTCCACGATGATGAG	1020

Db 1021 AGAGTATCATTAACAGCATCATTGGATTCCACGGTCAGAGTGTGGATGTTAATAACAGGT 108

Db	1081	GAATGCTAAACAGTTCATTCACCATTTGTAGACAGTTCGCCTTGCTTCAATTAAT	114
OY	1141	ggcattgagtgacctgtctccaagaatcgtttccattgctgtataggaatatagccctccca	120
Db	1141	GGCATTGAGTGACCTGCTCCAAAGATCGTTCATTTCTATGCGATATGGCTCCCA	120
OY	1201	actgacattaccctccggaagtgatctgtcgacaacgagctctgtcattgtcttgaagc	126
Db	1201	ACTGACATTACCCTCCGAGGGTCTGTGTGACACCGACTCTGTCATGTTGTAGAC	126
OY	1261	tttgatgacaagatacatgtttcttcgcatactgaggatagaaactlaaaagatacgaacaa	132
Db	1261	TTTGATGACAAGTACATTGTTTTCGATCTGGGGATGTGAACCTATTAAGCTATGGAACACA	132
OY	1321	agtaactbtgaaattctgaagaccttaattggacacaaacgaggaattgtcgttttgcaag	138
Db	1321	AGTACTGTGAAATTTGTGAAGSACCTTAAATGGACACAAAGAGGCAATTGCGCTTTTGCAg	138

```

|||||
Db 1381 TACAGGACAGCGTGTAGTGAAGTGCCTCATCTGCACACATCATGATTAAGGACATTA 1440
1441 gaatggtgcatgtttactagatgttagaagccatgaagaattggtgtgtatctga 1500
1441 GAATGTGGTGCATGTTTACGAGTGTAGAGAGCCATGAGGAATTTGGTGGTGTATGCA 1500
1501 ttgtataaagaagaatagtcagttgggacctatgataaataaagtgtggagctt 1560
1501 TTGTATACAGAGAGTATGATGAGTGGGCTTATGATGAGAAAATTAAGTGTGGATCTT 1560
1561 gtagcgtcttgagcccccgtctctcagaggaacactgtctcagaccccttgtag 1620
1561 GTGGTGTCTTGGACCCCGTGTCTCTGAGGAGACCTGTGTACGGACCTTGTGGAG 1620
1621 catccggaagatttctcgaactagttgataatccagattgtcagatgacat 1680
1621 CATTCGGGAAGATTTTGTGACAGTGTGATGATTCAGATTCAGTGTGATTCACAT 1680
1681 gatgaacaatctcatctcagactctcctaataatgatccagctgcccgaagctga 1740
1681 GATGACACAAATCCTCATCTGGGACTTCCCTTAATGATCCAGCTGCCCAAGCTGA 1740
1741 cgtccctctcgaacatacacacatcctcagataataaacaatacactgacctcat 1800
1741 CGTCCCTCTCTGACATACACATGACATCTCCAGATTAATTAACATACACTGACCTCAT 1800
1801 acttgcgaagaccataaagtgcgtatattaacgtatcgcgaataccagagatgag 1860
1801 ACTTGGCCAGGAGCCCATTAAGTGTGGGATTTAAAGTATCTGCCAATACGAGATG 1860
1861 aacaagaatgaacaataaactactgcccagttccctcgtgactagccgagaagaag 1920
1861 AACAGACGATACATTAACACTACTGCCAGTTTCCCTGGAATGACCGAGAGAGAGG 1920
1921 tgaactcctgttggagacacagttgtctcagctcggccaggaacggtctactca 1980
1921 TGAGACTCCTGTTGGGACACAGTTGGTGTGCACTGGCCAGAGCGTCTACTCACACA 1980
1981 actgactgtctcagatgtctatcacaagaatgtctctatcaattgtgaatgtgaac 2040
1981 ACTGACTGTCTCAGTGTCTGTATCAGAAAGATGCTTCTATCAATTTGTGAATTTG 2040
2041 tttaaacctccctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2100
2041 TTTTAAACCTCCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
2101 cagacaaagtgactataataatattagttgtttgcagaaaaaataa 2151
2101 CAGACAAAGTGACTATATAATATTATTAGTTTGGCCAGAAAATAAATAA 2151

RESULT 3
HSBTRCP 2151 bp mRNA linear PRI 07-JUL-1998
LOCUS HSBTRCP
DEFINITION Homo sapiens mRNA for beta-transducin repeat containing protein.
ACCESSION Y14153
VERSION Y14153.1 GI:2995193
KEYWORDS beta-transducin repeats; beta-TRCP gene; WD repeat.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2151)
AUTHORS Margolin, F., Bour, S.P., Durand, H., Selly, L., Benichou, S.,
Richard, V., Thomas, D., Strebel, K., and Benarous, R.
A novel human WD protein, h-beta trcp, that interacts with HIV-1
Vpu connects CD4 to the ER degradation pathway through an F-box
motif
JOURNAL Mol. Cell 1 (4), 565-574 (1998)
MEDLINE 98325370
REFERENCE 2 (bases 1 to 2151)
AUTHORS Benarous, R.

```

```

TITLE Direct Submission
JOURNAL Submitted (03-JUL-1997) R. Benarous, INSRM - I.C.G.M., Laboratoire
Interactions Proteiques, CHU Cochin, 24 rue de Fg-St-Jacques, 75014
Paris, FRANCE
FEATURES
Source 1. 2151
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Jurkat T-cell"
/cell_type="T lymphocyte"
/tissue_type="lymphoid"
70..1779
/gene="beta-TRCP"
70..1779
/codon_start=1
/product="beta-transducin repeats containing protein"
/protein_id="CAA74572.1"
/db_xref="GI:2995194"
/translation="MDPAEAVLOEALKFMNSEREDNNGPPIKIPKNSLROTY
NSCARLCLNDIETVCLASTYAKTKENYAKKRLANGTSMVPPROKLSAYEKELIV
KYEQWSESDVEFEVHELISOMCHVGHINSLKPLDORDITLALPARGDLHENT
LSYLDAKSLCAAEVCKEMVRYVSDMLMKLIERVVRDSDLMRGLAERGGQYLFK
NKPDPNAPNPSFYRALYPRKIIDIEPTIESNRCGHSLSQRIHCRSETSKYCLQYD
DOKIVSGLRDNTIKIDKNTLECKRLITGTGSLVCLQYDERVITIGSSDSYRWAD
NTGBMINTLIIHCEAVLHLRENNGMATVCSKDSILVMDMASPTDITLRLVLRHAA
VNVVDPDKYIVASGDRITKYMNSTCEFPVRLNHRKRGICLOVRODLVYSSDN
TIRLMDIEGCACLRVLEGEHELVRCLRPDKRIVSAYGKIKYMDIVALDPPADG
TILRLTVHNSGVHFRLODFEOLVSSHDITLIMDLNPAQAEPPRSRYITTY
ISR"
repeat_region
259..544
/rpl_family="WD"
/rpl_unit=516..544
BASE COUNT 628 a 467 c 513 g 543 t
ORIGIN
Query Match 100.0%; Score 2151; DB 9; Length 2151;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 tgcgttgctgctggcctcggccaccaaaggcggccggcggaagacggaacgagtgcc 60
1 TGCGTGGCTCGCGCTCGGCACCAAGGCGGCGCCCGCGGAGAGCGACAGTGGCC 60
61 tcgggatataggaccggcgcgagggcggtgtgtaagaagaaggaactcaagttatgaat 120
61 TCGGGAATTATGAGACCGCGCGAGGCGGTGTGCAAGAGAGGCACTCAAGTTTATGAAT 120
121 tccctagaagagaagactgtaataatggcgaaacccctcctagaagaataaccagaagaag 180
121 TCCTCAGAGAGAGAAAGACTGTAATAATGCGCAACCCCTAGGAAGATATACAGAGAG 180
181 aattcaacttagaagaacataacacagctgtgtccagactctgtcttaaaccaagaagaagta 240
181 AATTCACTTAGACAGACTATACACAGCTGTGCCAGCTTCGCTTAACCAAGAAAGCGTA 240
241 tgtttagaagacactgtcctatgaagaactgtagaattgtgtggccaaacaaacttgccaat 300
241 TGTTTAGACACACTGCTATACAGCTGAGAAATTTGTGTGGCCAAAACAAACTTGCCAAT 300
301 ggcacttcagatgatgtgtgcccagaacgaacgaacactcgaagaactatgaagaagaa 360
301 GGCACTTCAGATGATGATGTGCCCAAGACGAAACTCTCAGCAACCTATGAAGAAAGAA 360
421 aagaactgtgtgtcaataactcttgagcagtggtgcagagtlcagatcaagtggaattgtgt 420
421 AAGGAACGTGTGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
421 gaacacttatataccaatgtgtcattataccaacatgtggcacaataaactcgtatctaa 480
421 GAACTTTATATCCCAATATGTGTATTATACCAACATGGGCACATTAACCTGTATCTTAA 480

```


Qy	1801	accgcccaggaaccattaaagttgggagttttaaagctatctgcgaataccagaatgac	1860
Db	1801	actggcccaggaaccattaaagttggcgattttaaagctatctgcgaataccagaatgac	1860
Qy	1861	aacacagctaaacaactcaactgcccaggtttccctgacgtacgcagaagacagagcgt	1920
Db	1861	aacacacagttacatcaactcaactgcccaggtttccctgacgtataccgcagagacagggcgt	1920
Qy	1921	tgaagctcctgttggacacagttggtctgcagtcgagccaggaacggtctactcagcaca	1980
Db	1921	tgaagctcctgttggagacagtggtgctgcagtcgagccaggaacggtctactcagcaca	1980
Qy	1981	actgactgcttcagtgctgctcatcagaagaatgcttctcatcaattgtgaagattggaac	2040
Db	1981	actgactccttcagtgctgctcatcagaagaatgcttctcatcaattgtgaagattggaac	2040
Qy	2041	tttaaacctcccccctcctcctcctcttcacactcgcacactagtttctccatggttc	2100
Db	2041	tttttaaacctcccccctcctcctcctcttcacactcgcacactagtttctccatggttc	2100
Qy	2101	cagacaagaagtactataataatattagtggttttggccagaaaaaaaaa	2151
Db	2101	cagacaagaagtactataataatatttagtggttttggccagaaaaaaaaa	2151

LOCUS	AF081887	2175 bp	mRNA	linear	ROD 07-OCT-1999
DEFINITION	AF081887	Mus musculus ubiquitin ligase FWD1 mRNA, complete cds.			
ACCESSION	AF081887				
VERSION	AF081887.1	GI:4336326			
KEYWORDS					
SOURCE		house mouse.			
ORGANISM		Mus musculus.			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS		1 (bases 1 to 2175) Hatakeyama,S., Kitagawa,M., Nakayama,K., Shirane,M., Matsumoto,M., Hattori,K., Higashi,H., Nakayama,K., Onoe,K., Good,R.A. and Nakayama,K.-I.			
TITLE		Ubiquitin-dependent degradation of IkappaBalpha is mediated by a ubiquitin ligase Skp1/Cul 1/F-box protein FWD1			
JOURNAL		Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3859-3863 (1999)			
MEDLINE		99199275			
PUBMED		10097128			
REFERENCE		2 (bases 1 to 2175) Hatakeyama,S. and Nakayama,K.-I.			
AUTHORS		Direct Submission			
TITLE		Submitted (04-AUG-1998) Department of Molecular and Cellular Biology, Medical Institute of Bioregulation, Kyushu University, 3-1-1, Maedashi, Higashi-ku, Fukuoka 812-8582, Japan			
JOURNAL		Location/Qualifiers			
FEATURES		1..2175			
source		/organism="Mus musculus"			
		/db_xref="taxon:10090"			
		91..1800			
		/note="similar to beta-transducin repeat-containing protein trpc; F-box protein; contains W40 repeats"			
		/codon_start=1			
		/product="ubiquitin ligase FWD1"			
		/protein_id="AADI7755.1"			
		/db_xref="GI:4336327"			
		/translation="MDPAAYIOEAKLFENNSPERDNGGEPPRIIPKENSRLORYTNSCARCIQMOETVCLSTAMKTENCYAKAKLANGSSITVYRKQKLSASRYKEELCVLSEFQESDQVEFVEHLISOMCHVOGHINSYLPMLORPITALPKGIDHIAENILSLPAKLSICAAELKAEKWRVTSQGMKTKIEWVFSDLSKGLARROWGYLKRINPEENAPNSFYKALPKIIDIETIESMRCGRHSIQHCKSETSKCVQYOVDMKIVGLRNDTIKIMDKSTLECKRLIGHGTSYVLCQDSTGVSMVMDGKIVSGNLTILHHCFAVILHFRFNNGMVCYSKRSIAYWDMASPDIPLRLVGLHRAAPAVNYVPDDKTYIYASGDRTIKVMNTSTCEPVTLLGHKRGTAICQYRDLVSGSDSTIKLWMEICGACLRVLEGEHELVRKRIENPKRIVSGAVDCKIKWMDLMAALDPAPAPAGTTICLRVLVHSSKVPRLQFDEFQIVSSSHDITLLWDLNDPAHAIEPFRSPRTITYT			
CDS					

	IS R ⁿ	
BASE COUNT	589 a	528 c
ORIGIN		
		567 g
		491 t

Query Match	78.28;	Score 1682.2;	DB 10;	Length 2175;
Best Local Similarity	88.58;	Pred. No. 0;		
Matches 1909; Conservative	0;	Mismatches 228;	Indels 20;	Gaps 7;

[illegible]

Db	271	GAACAGATATGTTAGCAGACAGCTGCATGAAGACTGAGAAATGTTGTGGCCAAACAAACAA	330
Oy	252	cttgcgaatgycactlcccaatgataatgctgccccagcaacggaactctcaagaagctat	351
Db	331	CTTGCCATGTGCACCTTCCACTATGATGTTGGCCCAACCAAGGAAACCTCTCACCAAGCAT	390
Oy	352	gaaaggaagaagagcctgtgtcacaatattgagcagtggtcagaagtcagaatcaagt	411
Db	391	GAAGAAGAAAAGAACCTGTGTGTCMAATACTTTGAGCAGTGGTCAGAGTCCAGATCAAGTG	450
Oy	412	gaattgylgaaacattatataccaatlygtcatttacaatayggacataaactc	471
Db	451	GAATTTGTGGAACTTTATATCCCAAAAGTGTGCATTACCAACATGGGACATAAACCTG	510
Oy	472	tatttcaactatggtgcagagaagatttcaaacgctctccagctcggagattgat	531
Db	511	TATCTTAAACCTATGTTGCCAGAGAGATTTCATACGTCTGTCCACGTGGGATTTGAT	570
Oy	532	catalgcctagaaacattctcgtatataccttggatgcgaatacactatgtcctctaact	591
Db	571	CATATGCCGTGAAACATTCTGTGTCATACCTGGAGGCCAAATCACTATGTGCTGTCACTT	630
Oy	592	gtlgtgcaagaatlygtatccagatgacctctgatatgcaatcgttggaaagaactatcag	651
Db	631	GTGTGCAAGAAATGTATCCAGATGACCTCTGATGGGATCTGTGAAAGAGCTTATACAG	690
Oy	652	agaaatgltcagagacagatctctcgttggagagccttgcagacagaagaagatggagacag	711
Db	691	AGAAATGTGCAGAGACAGTCTGTGTGTGGAGAGGCTGTGGCAACGAAGAGATGGGGACAG	750
Oy	712	tatttattcaaaaaaacactcctctgcagcgggaatlytctctccaactctttatagaaca	771
Db	751	TATTTATTTCAAAAACAAACCTCCTGTGACGGGAATGTCTCCCAACTTTTATATAGACA	810
Oy	772	cttatcciaaaattatatacagacatttggacaataagatctcaatttggagtgtggaaaga	831
Db	811	CTTTATTCCTAAATTTATACAGACATTGAGACAAATGAAATCTAAATTTGGAGATGTGGAAGA	870
Oy	832	catatgttllacagaagatltcaatcgtccgaagtggaaaacagaagaagatllactgttllacag	891
Db	871	CATGTGTTACAGAGAAATTCCTACTGCCCAAGGTGAACAGCAAGAGATTACTGTTTACAG	930
Oy	892	tatagatataagaanaataagtaagcggccttcttgagaataacacataaagatctggatataa	951
Db	931	TATATATATAGAAATAGTAATAGTAAAGCGGCTTCGAGACAAACATCAATCAATCTGGGATTA	990
Oy	952	aacacattggaatgcagaagcgaatltcccaoaagccatacaggttcaagctcctcgtctccag	1011
Db	991	AACACATTTGAAATGCAAGCAATTTCCAGAGGCGATACAGGTTCAAGTCTCTGTCTCCAG	1050
Oy	1012	tatgatagaagatgatacataaagaatcatatggatltccagcgttcaagagtgtggatgta	1071
Db	1051	TATATATAGAGAGTGAATCAAAACAGGATATATCGATTCACCGTCAAGAGTGTGGGATGTA	1110
Oy	1072	aatacaggtlaaaatgctlaaacaacgttgcatacacaatttggaaagcagcttgcacttgcgt	1131
Db	1111	AATACAGGTGAATATGCTAAACACAGTGTATTCACCAATTTGGAAGCAGTCTCGACTTGGCT	1170
Oy	1132	ttaataataygcatagtatggtgacctgcctccaagaatcglttccattgctgtatgagata	1191
Db	1171	TTTCATATATATGGCATGTATGGTGAACCTCTCCAAAGATTCGTTCCATTGCTGATTTGGGATATG	1230
Oy	1192	ggcccccacaactacattaccctccgagaggtgctcgttgcagacacagagctgcgttcaat	1251
Db	1231	GCCCTCCCAACTACATTTACCTCCGGAAGGTTGCTGTGTGGACACCGAGCTGTCTCAT	1290
Oy	1252	gttgtagacttggatgacaagatacatlgttcttgcatacttgggataagataaactataaagta	1311
Db	1291	GTTGTGAACCTTTGATACAAAGTCAATTTGTTCTGCATCTGGGGATAGAACTATAAAGSTA	1350
Oy	1312	ttggaacacaagtaactgtgtaatttggaaagaccttaaatgtaacaaacagcgcatggcc	1371
Db	1351	TGGAAACAAAGTACTGTGTGAATTTGTGAAGACCTTAAATGAGACAAACGAGGAGATTGGCC	1410

OY	1372	tgttttcagataagggaacggctgtagtgatgagctcatcgcacaactatcgattta	1431
OY	1411	tgttttgcagtacagggaacggctgtagtgatgagctcatcgcacaactatcgattta	1470
OY	1432	tggagacataagaatcgtgctgacatgcttaccagagtgcttagaaagsgccatgaaatgtgcgt	1491
Db	1471	TGGAGCATATGAATGTGCTGCATGTTTACGACTCTTAGAAAGCCATGAGGAATTGGTGCGT	1530
OY	1492	tgtatcgaatttgtatacaagaagatagtcagtgaggccatgatgtgaaaaattaagt	1551
Db	1511	TGTATTCGATTTGGATAACAAGAAGATAGTCATGGGGCCTATGATGGAAAAATTAAAGTG	1590
OY	1552	tggagatctgtgtgcgcttttgagcccccgctgtccctcgaaggagcaactgtgtaccagacc	1611
Db	1591	TGGAGTCTTGTGTGCGCTTGTGAGCCCCGCTGTCTCTGCAGSAGCACTGTGTACGAGACC	1650
OY	1612	ctgttgagagcatcccggaagagattlltcgactaacagtttgatlgbaatlccagatgtcagt	1671
Db	1651	CTTGTGGAGCAATCCGGAAGAGTATTTTTCGACACAGTTTGGATGGAATTCACGATTTGTCAGT	1710
OY	1672	agttcacatgatagacacaatccatctctcggagcttcccataatgacctcgtgccaaact	1731
Db	1711	ACTTACATGATGACACACATCCTCATCTTGGAGCTTCTTAATGATCCAGCTGCCAAGCT	1770
OY	1732	gaaccccccgyltccccctlctcgaacatacacactacatccccaataa	1779
Db	1771	GAAACCCCCCGCTTCCCCTTCGACACATACACTCATCTCCAGATTA	1818

RESULT	7
LOCUS	BC003989
DEFINITION	Mus musculus, beta-transducin repeat containing protein, clone MGC:7517 IMAGE:3491843, mRNA, complete cds.
ACCESSION	BC003989
VERSION	BC003989.1 GI:13278339
KEYWORDS	MG.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	1 (bases to 2081)
TITLE	Strausberg,R.
JOURNAL	Direct Submission
REMARK	Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Gilbert Smith, Ph.D. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Sequencing group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcdgapxill.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAP Plate: 8 Row: 0 Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6753209.
Location:Qualifiers

FEATURES
source

CDS			/tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."		
			/clone_lib="NCI_CGAP_Lu29"		
			/lab_host="DH10B"		
			/note="Vector: PCMV-SPORT6"		
			7..1716		
			/codon_start=1		
			/product="beta-transducin repeat containing protein"		
			/protein_id="AA03989.1"		
			/db_xref="GI:13278340"		
			/transl_table="MDPAVAVLQERALKPMNSSEEDCNNGEPKILPEKNSLRQYNSCARLIMQETVCLSTAMKTENCARAKLANGTSMIVPEKORLSAYEKELCYKPEWMSSEDOVEFEVHLISOMCHYOHGINSYKPMLOPFITALPARGDHAENI LSYDAKSICAAELVCKEMRYRYSQGMIMKILPEWYRPSDILMRLARGMGCIYKKNKPPENAPNSFYRLPKIITODIETIESNRCRSHLORHCHSETSKVYCIOTD DOKIVSLKDNITKIMDSSTLECKRILITGHGVSVCIDYDEVVITIGSSDSTVRADY NAGEMLNTLIHCEAVLHLRFNNGMVMVCSKDRSIAVDMASPDITLRVLVGRRA VVWVEDDKYIVSASGDRITKYWNTSTCEFTVNGHGRIGIACIQDRLYVSGSDN TIRLMDIEGACLRVIEGHEELRCIRIDNKRIVSGAYDGIKIVMDLMAALDPRAPAG TICTLRTIVHSGRVFRLOFDEFOIYSSSHDDIILIMDFLNDPAAHAERPRSPRTYTT ISR"		
BASE COUNT			568 a	502 c	529 g
ORIGIN			482 t		
Query Match			75.8%; Score 1630.4; DB 10; Length 2081;		
Best Local Similarity			88.5%; Pred. No. 0;		
Matches 1853;			Conservative 0; Mismatches 221; Indels 20; Gaps 7;		
QY	64	gacattatgaccgagcgatgctgctgcaagaagaagcaactcaagttatgaatcc	123		
DB	1	GCATTATGACCCGCGACAGCGGCTGTCAGAGAGAAACCGCTTATGAAATGCC	60		
QY	124	tcaagagagaagaacgctgataatgagcaaccccttagaagaataatcagaagaagt	183		
DB	61	TCAGAGAGAGAAAGCTGTATATGCGCAACCCCTAGAGAGATATACCGAGAGAAAT	120		
QY	184	tcaactagacagacatacaacaagctgccaagactgcttaaaccaagaagaactatg	243		
DB	121	TCACCTTAGACAGACTTACAAACAGCTGTGCGAGGCTTTGCATTAACCAAGACAGATATGT	180		
QY	244	ttagcaagaacgctgctatgaagactgagaattgctgagccaaagaactgccaatggc	303		
DB	181	CTAACCAAGCACTGCTATGAAAGACTGAAATTTGTGCGCAAAAGCCAACTTGCCAAATGCG	240		
QY	304	acttccaagatgattgctgccaagaagaactctgcaagaactaataaagaagaagaag	363		
DB	241	ACTTCCAGCATGATTGTGCCAAGACGCGAAACTCTCAAGCAAGCTATGAGAGAGAAAG	300		
QY	364	gaactgtgttcaaatactttagcagctgctcagaagctcagaatcaagtggaaattgtgaa	423		
DB	301	GACTGTGTGTCAAGTATTTTGAAGAGTGTGTCAGATCTGATCACTGGCAATTTTGAGAA	360		
QY	424	catcttatatcccaaatgtgtcaatcccaaatgagacataaactcgtatcttaacct	483		
DB	361	CACCTTATATCCCAATATGTGTCTACACAGCATGGGACATCAACCTTCACTTAAACCT	420		
QY	484	atgttccaagaagattcaactgctgccaagcagcaggaatttgatatataatcgtctgag	543		
DB	421	ATGCTGCAGAGGATTTCACTAATGCACTGCCAGCACGGGGTCTGGACACACATCGCTGAG	480		
QY	544	aaatctgtcatactcgtgcaaatcaactatgtctgtcgtgaacttgtgtgcaagaa	603		
DB	481	AACATTTCTGATCTTGTGAGCGCCAAAGTCACTGTGTGCTGAGCTCGTGAGCAAGAA	540		
QY	604	tgtgtacgaatgacccctgctgcaatgctgtggaagaagcttatcgaagaagaatggtcag	663		
DB	541	TGGTACCGCGTCAAGTCGACGCGCATGCTGTGAAAGAGCTCATCCAGAGAGATGTCAGG	600		
QY	664	acagattctctgtggaagccctgcaagaagaagagttggaagacattatcaca	723		
DB	601	ACGACTCTGTGTGGAGAGGCTGGCAGAGAGCGGCTGGGAGACACTTATTTCAAA	660		
QY	724	aaaaactctctgacggaatgctctccccaactctttatagagacttactcctaa	783		
DB	661	AACAAACCTCTGTATGAGAGCGCTCTCCCAACTCTTTATAGAGGCTTTATCCAAA	720		
QY	784	attataagaacattgagacataagaatctaatgtgagatgtggaagacataattacag	843		
DB	721	ATCATACAGACATTTGAGACATAGAGTCCAAATTTGAGAGATGTGGCGACATAGTTTACG	780		
QY	844	agaattcagcgcgaagtgaaagaagaagaagattactggtttacagttatgatgacag	903		
DB	781	AGAAATCCACTCCCGAGTGAACAAAGTAAAGGGGTTTACGTTTACAGTACGACGACGACG	840		
QY	904	aaaatgaagaagcccttcgagacaacatacagaatctcgggaataaacaacatgaa	963		
DB	841	AAGATAGTCAGCGGCTTCGAGACACACATCAAGATCTGGATTAACACACATCGGA	900		
QY	964	tgcgaagcaattctcaacagcctacagaagttcagttcctgtctcagatagttagaga	1023		
DB	901	TGCAAGCGGATTTACAGCGGCGCACAGGGGCTCCGCTGTGTGTCAGTACGATGAGAGG	960		
QY	1024	gtgtatcaataaagatcatcggatccagcgtgaagtggtggaatgtaataagatgaa	1083		
DB	961	GTGATCATCACAGGCTCTCTCAGACTCCAGCTCAGAGTGTGGATGTAATGCAAGGTGAG	1020		
QY	1084	atgtctaaacacgttgatcaccattgtgaaagcaattctgcaattgctgttcaataatggc	1143		
DB	1021	ATGCTAAACATTTGATTCACACTGAGAGCGCTTGTCGACCTGCGCTTCAATTAATAGCG	1080		
QY	1144	atgatgtgacgctgctccaagaatcgcttccatgctgtatgaggaatagagccctcccaact	1203		
DB	1081	ATGATGGTGACCTCTCCAAAGACCGTTCCATCGCTGTGGGATGATGGCTTCCCAACT	1140		
QY	1204	gacattaccctcccgagaggtgctgtgctggaacacagctgctgcaatgtgtgaactt	1263		
DB	1141	GACATCACCTTCAGAGAGGCTCTGTGGAGACCGGACCTGGGTCATGTTTGAAGATTT	1200		
QY	1264	gatgacaagtaactgtttctgcatctgaggaatagaactaaagataggaacaagt	1323		
DB	1201	GATGACAGTATCATGCTTCTGCTGTGAGATGAGATCAATTAAGGTGTGGAACACAAAT	1260		
QY	1324	actgtgtaattgtgaagacacttaaatggaacaaagaagcctgtgtgtgtgagttac	1383		
DB	1261	ACCTGTGATTTCTGAAGGACCTTAATGTGGGCAAGGCTGGCATCGCTGTGGCAATAC	1320		
QY	1384	agggaagcgtgtagtgaagtgctcatctgacacacatacagattatgagacataaa	1443		
DB	1321	AGAGACAGGCTGTGTGTGAGGCGCTCCTGTGACAAACACATCAGCGTGGGACATAGAG	1380		
QY	1444	tgtgtgcatgtttaagagtgtagaagccatgagaagaattggtgctgtatcgaatt	1503		
DB	1381	TGTGAGCATCGCTGCGAGTGTGTGAGGCGCATGAGAGTGTGCTGACCTCATTCGATTT	1440		
QY	1504	gataacaagaagatagtagcagtgaggcctatgatatgaaataaagtgtgagctgtgtg	1563		
DB	1441	GATAACAAAGATGATGAGCGGAGCCATGATGTGGAATTAAGTGTGGAGATCTTATG	1500		
QY	1564	gctgcttggagcccgctgctcctgcaaggaacatctgtctacgagacccctgtggaagcat	1623		
DB	1501	GCTGCTTGGAGCCGCGTGCCTCAGAGGAGCACTGTGTGTGCGGACACTTGTGGAGCAT	1560		
QY	1624	tccggaagagtttctcgaactcagttgagaattccaagattgcagtagttcaatgat	1683		
DB	1561	TCTGGAAGATTTTCCGCTCTCCAGTTTGATGAATTCACAGATTGTCATAGTTTCACATGAT	1620		
QY	1684	gacacatccatctgagacttcttaaatgatccagctcccaagctggaaccccccgt	1743		
DB	1621	GACCAATTTCAATCTGTGGACTTCTCTGATGATCTCACTGTCTACCGCTGAACCGCCCGC	1680		
QY	1744	tcccctctcgaacatacacactatcatctccagataaataaacacatactgacctcact	1803		
DB	1681	TCCTCTTGTGGACATACACTACATCTCCAGATTAATAAACCCAAACACTGGCCTCATTAAT	1740		

QY 1210 accctccgaggggtgctgctgcagaccagctgctcgaatgctgtaacttgcac 1269
 DB 1141 ACCCTCAGAGGGGTGCTGCTGGAGACCGAGCTGGCGTCAATGTTGATGCTTATGAC 1200
 QY 1270 aagtaactgttctgcacatcctggatagactaaagatagaaacaactgtgt 1329
 DB 1201 AAGTACATGTTGCTCCCTGGAGATGACAACTAAAGTGTGGACCAAGTACCTGCT 1260
 QY 1330 gaattgttaaggaccttaattgacacaaagagcattgctgttggatagcaggac 1389
 DB 1261 GAATTGTGTAAGACCTTAATGGCCACAGCTGGCATGGCTGTTTGACATGAGAGAC 1320
 QY 1390 aagctgttagtagtgcacatcctgacaaactacatcaatltatggagacataagtgct 1449
 DB 1321 AGGCTGTGGTGGAGCGGCTCCTCTGACAAACACATCAGGCTGTGGACATAGAGTGGGA 1380
 QY 1450 gcatgttaccagagtgcttagaaggccatgaggaattgtgtgctgttattcgaattgataac 1509
 DB 1381 GCATGCTGCTCGAGTGTGGAGGGCCATGAGAGTGTGACGCTGCATTGATTTGATTAAC 1440
 QY 1510 aagaagatagctagtgaggccatgatgatgaaaaaattaaagtgtgagctgtgtgtgt 1569
 DB 1441 AAAAGATAGTGAAGCGGAGCGCTATGATGGAATAATTAAAGTGTGGATCTTATGGCTGCT 1500
 QY 1570 ttggaaccccgctgctcctgcagaggaactctgtctacaggaacctgtgtgagcaatccgga 1629
 DB 1501 TTGAGACCCCGTGTCTCCACAGGAGACTCTGTCTGCGGACACTGTGGAGCATCTGTGGA 1560
 QY 1630 aagaatttccagctacagctgttgatgaatcgaatctgtcgaagtgtcacaatgagaca 1689
 DB 1561 AAGATTGCTCGCTCCAGTGTGATGAATTTCAAGATTGTGACATGATGATGACACA 1620
 QY 1690 atcccatctggagactcctcctaattgatccagctgtcccaagctggaaccccccgtgtccct 1749
 DB 1621 ATTTCATGCTGGAGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1750 tctggaacatatacctacatctccagataaataaacatatacctacatctacatctccca 1809
 DB 1681 TCTCGACATATACACTACATCTCCAGATTAATAACCAACCACTGCGCTCATTAATTGCCCA 1740
 QY 1810 ggaaccatataagctgtggtatlaagatgacatgacatgacatgacatgacatgacatgacat 1869
 DB 1741 GGAATCTGTTAATGTTGCAATTTTAACAGACTGCAAGACCTGCAAGACCTGCAATTAAC 1791
 QY 1870 aacaaatcaactactgctcagctgtccctgt--gactagccagagagagagctgtgagact 1927
 DB 1792 AACATVCAAACTCTTACCCGGATTCCTCGACGAGATGAGCAGAGAGAGAGAGCTTTGAGACT 1851
 QY 1928 cctgtgtgagacagctgtgctgacgtcggccagagagagctgtcactacagacaactgact 1987
 DB 1852 CCTGTGGGAGACACAGCTGCTGAGAG--CCGACCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1907
 QY 1988 gcttaagctgtgctacgaagatgcttctatacaatgtaagatgtaagatgtaagatgtaagat 2047
 DB 1908 GCTTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1966
 QY 2048 cctccctcc 2057
 DB 1967 CTTCCCTCC 1976

RESULT 9
 AF099932 1712 bp mRNA linear ROD 21-JAN-1999
 LOCUS AF099932
 DEFINITION Mus musculus beta-Trop protein E3RS-Ikappab mRNA, complete cds.
 ACCESSION AF099932
 VERSION AF099932.1 GI:4008019
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1712)
 AUTHORS Varon, A., Hatzubai, A., Davis, M., Lavon, I., Amit, S., Manning, A.M.,
 Andersen, J.S., Mann, M., Mercurio, F. and Ben-Neriah, Y.
 TITLE Identification of the receptor component of the
 JOURNAL Nature 396 (6711), 590-594 (1998)
 MEDLINE 99075339
 REFERENCE 2 (bases 1 to 1712)
 AUTHORS Varon, A., Hatzubai, A., Mercurio, F., Manning, A.M., Andersen, J.S.,
 Mann, M. and Ben-Neriah, Y.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-1998) Immunology, Hebrew University of Jerusalem,
 Bin Karem, Jerusalem 91120, Israel
 FEATURES
 source
 1. 1712
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1. 1710
 /note="beta-transducin repeat-containing protein; receptor
 component of Ikba-ubiquitin ligase"
 /codon_start=1
 /product="beta-Trop protein E3RS-Ikappab"
 /protein_id="A00870.1"
 /db_xref="GI:4008020"
 /translation="MDPAEAVLQERALKFMNSEREDCNCEPPRKIIIPKNSLRQY
 NSCARCINQETVCLSTAMKTECNVAKKLNKSSMIVPORKLSAYSEKELCY
 KYEBOESDQVEPEVHLISOMCHYOHGINSYLPMLORDITLAPARGDHAENI
 LSYLAKSLCAELVCKEYRYRTSGMKLRIEMVNTDSIMRLARRRGQTLFK
 KRPPDNAPNSFRALYRKIIIDLETIESMNCGRHSIORHCSESKVCLQYD
 DKIVSGLDNFTIKIWDKSTLECKRIIVGHSVLCIDYDERVITGSSDSYRVMDY
 NGENMLNTLIHCEAVLIHLRFNNGMNVCSKDRSLAVMDMSPDITRLRYVGHRA
 VVWVPEDDKYIVASASDRTIKVNTSTCEFEYTLNGHKGIALQYRDLRYVSGSDN
 TJLMDIECGACLRYLEGHLELVRICRFENKRIIVSGAVDKIKVMDLMAALIDPPAPAG
 TICTRLIVSHSRVRLQDEHQIYSSSHDITILIMDLNDPMAIAEPRPSRTYTY
 ISR"

BASE COUNT 469 a 399 c 453 g 391 t
 ORIGIN

Query Match 67.5%; Score 1451.2; DB 10; Length 1712;
 Best Local Similarity 90.5%; Pred. No. 0;
 Matches 1549; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 70 atgagccggcgagggcggtgctgcagagaaagagcactcaattatgtaattccctcagag 129
 DB 1 ATGAGCCGGCGGCGAGGGGCGTGTGCGAGGAAGAGCCTTAAGTTATGATTCCTCAGAG 60
 QY 130 agagaagactgttaataatgagcgaacccctcagagaataatccagagaagaactcact 189
 DB 61 AGAGAAGACTGTATATATGCGGAACCCCTAGAGATTAATACAGAGAAAGATTCACAT 120
 QY 190 agacagacatacaaaagcgtgtgcccagactcgtcttaaccaagaagaacatgatttga 249
 DB 121 AGACAGACTTACAAACGCTGTGCGAGGCTTTGATTAACCAAGACAGATAGTCTAAC 180
 QY 250 agactgtctatgaagactgaagtgtgtgcccacaaacaaactgtccatgtccactcc 309
 DB 181 AGCACTGCTATGAAGACTGAAATGTGTGCGCAAGCCAAACTGTCCTCAATGCGACTTC 240
 QY 310 agtatgtgtgtgcccagaagcgaaggaactcctcagcagactatgaaaggaagaaggaact 369
 DB 241 AGCATATTTGTGCGCAAGAGCGGAAACTCTCAGCAAGGTATGAGAAAGAAAGAGCTG 300
 QY 370 tgtgtcaaatctttagagcagtggtcagagtcagatcagatggaattgttgaacatctt 429
 DB 301 TGTGTCAAGATTTTGAAGCAGAGTGCTGATCAAGTGAATTTGTAGAACACCTT 360
 QY 430 atatcccaaatgtgtcattacaacatgagcacaataaactgtatcttaacactatgtgt 489
 DB 361 ATATCCCAATGTGTCACTACAGATGCGGACATCAATCACTCTACTTAAACCTATGCTG 420
 QY 490 cagaagatttataactgtctgtcagactcgtgggattgagatcatatgctgtgagaacatt 549

QY 130 agagaagactgtaataatgagaaaccccttagaagaataataccagagaagaattccact 189
 Db 61 AGAGAAAGCTGTATATATATGCGCAACCCCTAGAGAAATATACGAGAGAAATTCACCTT 120
 QY 190 agacaacatacacaagcgtgcccagactctgtttaacaagaagaacagatggtttagca 249
 Db 121 AGACACACTTACACAGCTGTGCGAGGCTTTGCATTAACACAGACAGACATATGCTTACA 180
 QY 250 agcaatgctatgaagaactgagaaattgltgccaanaaactgccaatgccaattcc 309
 Db 181 AGACAGCTATGAGACTGATAAATGTGTGGCCAAAGCCAAACTTGCCAAATGGCACTTCC 240
 QY 310 agtaattgtgcccagaagaaacactctgaagaactatgaataaagaagaagaactg 369
 Db 241 AGCATATTGTGCCCCAGAGAGCGGAACCTCTAGCAAGCTATGAGAAAGAAAGAGAGCTG 300
 QY 370 tglgtcaaatcttgaagcagtgctcaagatcagaatgaatgaatgaaattgtggaacatct 429
 Db 301 TGTGTCAAGTATTTGAGCAGTGTGAGAGTGTGATCAAGTGTGAATTTGTAGAACACCTT 360
 QY 430 atataccaaatgtgtattaccacaatgagacataaactgtatcttaaacatctgtg 489
 Db 361 AATATCCCAATGTGTACTACACAGATGGGCAATCACTCTACCTAAACCTATGGCTG 420
 QY 490 cagaagatltcataactgtctgcagctcgaggatltgacatcatcgctgagaacatt 549
 Db 421 CAGAGGATTTCAATACCTGCATCGCAGAGAGCGGGTGTGACACATCGCTAGAACATTT 480
 QY 550 cgtctacactgtagagcaaatcaatcagctgctgcaactgtgtgcaagaaatgtgac 609
 Db 481 CGTGTACTTGTGACCCCAAGTCACTGTGTGCTGTGATGCTGTGAGAAAGAAATGTGAC 540
 QY 610 cgaatgaactctgtagagcagctgtgtggaagaagcttatcagaagaatgtgagaacatt 669
 Db 541 CCGGTGACCTCGGAGCGCATGCTGTGTGAAAAAGCTCATCGAAGAGATGTGTGAGAGAC 600
 QY 670 tctctgtgagagagccctgagcaagaagaagatgtggagacatltatltcaanaacaa 729
 Db 601 TCTCTGTGGCAGGCGCTGCGACAGCGAGAGGCTGGGCGACATCTTATTTAAAAACAA 660
 QY 730 cctctcgaaggaatgctctcccaactctttttagagacattatccataaattata 789
 Db 661 CCTCTGATGAGACGCTCTCCCAATCTTTTATAGAGCCCTTTATCTTAAATTCATA 720
 QY 790 caagaactatgagacaataagatctaatgtgagatgtggaagacatagtttaagaagaatt 849
 Db 721 CAAGACATTTGAGACATATAGTCCCAATTGAGATGTGGCGCACATAGTTTACAGAGATTC 780
 QY 850 caatgcgaagtgaaacaaagaagaagtttacgttttaacagatgatgacagaata 909
 Db 781 CACTGCGGAGTGAACAAGTAAAGGGGTTTACTGTTTACATAGACAGACCGAGAAAGTAA 840
 QY 910 gtaagcgacctcgagacacacaataagaatctgagataaagaacattgagaagaag 969
 Db 841 GTACAGCGGCTTCGAGACACACCATCAAGATCTGGGATTAAGACACACTGGAAATGCAAG 900
 QY 970 cgaattctcacaagccatacagatgctcactgtctcgcagatgagagagatgtac 1029
 Db 901 CGGATTTCTACAGGGCCACACGCGGCTCCGCTGTGTGTGACGTAGCGAGGAGGTATTC 960
 QY 1030 ataaacagatcatcgatltccacggtcagatggtgtgagatgtaataagaagtgaaatgcta 1089
 Db 961 ATCAACAGGCTCTCTCAGACTCCACCTCAGAGTGTGGATGTGAAAGCAGGTGAGATGCTA 1020
 QY 1090 aacacgttgaattcacacatgtgaagcagatgtgcacgtgctgttcaataatgagatgag 1149
 Db 1021 AACACATTTGATTCACCACTGTGAAGCCTTTCGACCTCCCTTCATTAATGACATGATG 1080
 QY 1150 gtgaactgtccaaagatcgttccatgtgctgtatgagatagtgacctcccaactgacat 1209
 Db 1081 GTGACCTGTTCAAAGACCGTTCATGCTGTGTGGATATGGCTTCCCAACTGACATTC 1140

QY 1210 accctcggaggggtgctgctcgcagaccagagctgctgttcaatgtttagactttagac 1269
 Db 1141 ACCCTCAGAGAGGGGTGCTGTGTGGACACCGAGCTGCGGTCAATTTGTAGACTTTGATGAC 1200
 QY 1270 aagfacattgttctgcatctgggagtagaagacataaaggtaagtaagacaagaacttctg 1329
 Db 1201 AAGTACATTCGTTTCTGCTCTGGAGATGAGAACCATTAAGGTGTGGAACCAATACCTCTG 1260
 QY 1330 gaattgtgaagaccttaattgagacacaagaaggaactgtccgttttcagtaagagac 1389
 Db 1261 GAATTCGTAAAGACCCCTAAATGAGGCAACGCTGGCATCGCTGTTTGTGACATGACAGAC 1320
 QY 1390 aggtgtgtagtgaagtgctcatctgagacacacatacagatgaatgagacataagatgtgt 1449
 Db 1321 AGGTGTGTGTGAGACGGCTCTCTGTGACACACCATCAAGCTGTGGACATAGAGTGTGA 1380
 QY 1450 gcatgtttagagatgttagaagccatgaaaggaatgtgtgctgtttagatcgaattgaaac 1509
 Db 1381 GCATGCTGCGAGAGTGTGAGGCGCATGAGAGATTGTTGATGCTGATTCGATTTGATTAAC 1440
 QY 1510 aagagatagtcagtgtaggagcctatgtagaataaataaagtgtgagatcttgtgtgct 1569
 Db 1441 AAAAGATGACTGAGCGGAGCTTATGATGGGAAATTAAGTGTGAGATCTTATGCTGCT 1500
 QY 1570 ttgagaccctgctcctcgcagagacacatctgtctacagacccctgtggaacatccgga 1629
 Db 1501 TTGAGACCGGGGTGCTCCAGAGAGGACTCTGTGTGCGGACATTTGTGAGCATTTGTGA 1560
 QY 1630 agagttttcgaactacagtttgaatgtccagatgtgcaatgttcaatgtgacaca 1689
 Db 1561 AGAATTTCCGCGCTCCAGTTGATGAAATTCAGATTTGATGCTGATGATGATGACACA 1620
 QY 1690 atctctgtgagacttccaaatgacagctgagccaaagtgaaccccccttccct 1749
 Db 1621 ATTCTCATCTGGGACTTCTGGAATGATCAAGCTGCTGACGCTTAACCCGCCCTCCCT 1680
 QY 1750 tctcgaacatacacatcatctccagataa 1779
 Db 1681 TCTCGACATACACTTCATCTCCAGATTA 1710

RESULT 11
 XELSTRCP
 LOCUS African clawed frog beta-Trip mRNA, complete cds.
 DEFINITION
 ACCESSION M98268.1 GI:295542
 VERSION
 KEYWORDS beta-transducin repeats.
 SOURCE Xenopus laevis (library: S. cerevisiae expression library of X.laevis oocytes) cDNA to mRNA.
 ORGANISM
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
 Xenopus laevis
 REFERENCE
 1 (bases 1 to 1671)
 AUTHORS Spevak, W., Keiper, B.D., Stratowa, C. and Castanon, M.J.
 TITLES Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein with b-transducin repeats
 JOURNAL Mol. Cell. Biol. 13 (8), 4953-4966 (1993)
 MEDLINE
 FEATURES
 source
 Location/Qualifiers
 1..1671
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /cell_type="oocyte"
 /tissue="lib-S. cerevisiae expression library of X.laevis oocytes"
 1..1671
 /gene="beta-Trip"
 48..1604
 /gene="beta-Trip"
 /codon_start=1

	source	1..2134	/organism="Homo sapiens"
		/db_xref="taxon:9606"	
		/chromosome="5"	
		/map="5q35.1"	
	gene	/tissue_type="fetal lung"	
		102..1628	
	CDS	/gene="BTRCP2"	
		102..1628	
		/gene="BTRCP2"	
		/codon_start=1	
		/product="F-box and WD-repeats protein beta-TROP2 isoform A"	
		/protein_id="BAAG2329.1"	
		/db_xref="GI:7209809"	
		/translation="MEPDSVLEEDTKTEILMTISNGNSVIYSVRKRPSEFGNYOKERDLCTIKYFDPMSEDOVEFEVHEFLISMCHYOHHMINKLIEKVPDMLQORDITPLPROGDHLINELLII SYLDARSLCAELCYCKEMORYISEGMIIMKLIEKVRDPLMKGLSERGWDTLFKNRPTDGPNPSFYRLSPKIIDIDETESNRCGRINLDIQRSSNSKVYCLDYDDKE IISLRDNRSIKIMWTSILECLKVLGHTGSYLQYDERIVTSSSDSYAVMWVNNGE EVLNTLIHNHDAVHLHFRESNLAWTCSDKSIAWDMASTDIYLRVLVGHAAVVWV VDPEDKYIVSASGDRTIKVMSTCEPEVRTINGRKRIACIQYDRDLVVSASSONTLR LMDLEGACLRILVEGHBEIVRCIRFDNRKIYSSAAYDKIKWIDLQAALDPAPASTLC LRTIVESGRVFRQLPFDEFOQLISSHDDTLIMDFLVNPESAUNETSPSKTYITYISR "	
BASE COUNT	566 a	477 c	515 g
ORIGIN		576 t	
Query Match	41.1%	Score 883.4	DB 9; Length 2134;
Best Local Similarity	74.8%	Pred. No. 4.5e-243;	
Matches 1122;	Conservative 0;	Mismatches 371;	Indels 6; Gaps 1;
Dn	298	aatggcacttcacgatgatgattgttgcccaagaacaagaaacctctcagcgaaglatgaaaag	357
Dn	153	AATGGACATCATCGTGATGCOTCTCCAGAAAGAGGCCATCAGAGAACAATATCAAATA	212
Oy	358	gaaaaggacacgtgtgtgtaaatlactlttagaacgtgtgtaagatcgatcaagtgaattc	417
Dn	213	GAAAAAGACTGTGTATTAAATTTTGAACAGTGCTGTGAATCGAGATCAAGTGAATAATT	272
Oy	418	gtggaacatctatatccccaatgytgtataaccacaatlbggcacataaactcgtatctt	477
Dn	273	GTGGAACATCTTATTTCACGAATGTCATTATACGATGGACATATTAACTCTTAACCTG	332
Oy	478	aaacctatgttcagagagatttcaataactcgtctgcagactcgaggatltgataatc	537
Dn	333	AAGCCCATGTTGACACGGGACTTTATTACCCTTTACACAGACGACGCTTGATGATCACATA	392
Oy	538	gctgaacaacttcgtctactcgtgaatgcacaatcaactcatgtgtgcgcgaactttgyc	597
Dn	393	GCAGAAAACATCTTTTCGTACTCTGATGATCCAGAGTCTCTGTGTGCGACGAGAGCTGATAGT	452
Oy	598	aaggaatgtaacgaagtgtacctctlgatlgacatgctcgtgtaagaagcttalcgagaagaly	657
Dn	453	AAAGATGCGACGAGTAGTATCTCAGAAAGAAATGCTTTGGAAGAAACGTGATTGAACGAATG	512
Oy	658	gtcaggacaagttctctgttgagagagcctctgcaagaacaagaagatlgugagcatatta	717
Dn	513	GTAGCACTGATCCCCTTATGGAAGAGACTTTTCAGAAAAGAGGGTGGGATCAGTACCTG	572
Oy	718	ticaaaaaacaactcctcgacgggaatgcctctcccaactcttltatagacattat	777
Dn	573	TTTTAAAMACACCCACAGATGCG-----CCTCCAATTATCTTTTATAGTCATTTATAC	626
Oy	778	cctaataatatacagaagacattgagacaatagaaatctaattgagaglytggagaacatagt	837
Dn	627	CCAAAGATTATCCAGGATATAGAGACTTATACAATCTTAACCTGGCGGTGGACACACACAC	686
Oy	838	ttaagaagaattcactgcccgaagtgaacaagaagaagagtttactgttlcagaatgpat	897
Dn	687	TTGCAAGAGGATTCAGTGCCTCTGAAAATAGTAAAGGTGCTTACTGTTTACAGTAGCAT	746

Oy	898	gagagagaaatagtaagcggccttcgaagacaacacaaatcaagatcctgggtaaaacaca	957
Db	747	GATGAAAAAATTTATCATGAGGCTACGAGATTAATTCTATTAAAGATATGGATTAACACAG	806
Oy	958	ttggaatgcgaagcgaattctcacagggccataaaggtltaagtcctctgctccacgatgat	1017
Db	807	CTGGAATGTTTGAAGCGTTTAACAGAGACACACAGGCTGCTGCTCTGTCAGATGTGAT	866
Oy	1018	gagagatgatcatcaacagatcatcggattccaaagtcgaagtggtggatgtaaatata	1077
Db	867	GAGGGTGTCAATTGTAACTGACTCTCTCAGATTCTACGGTACGGGAAGTGTGGATGTGAACAG	926
Oy	1078	ggttgaatgcctaaacacgcttgattcaacctgttgaaagcagtcctgcacatcgcttcaat	1137
Db	927	GGTGAATGTTTAAACATTTGATTCACACCAATGAGGCTGTGATTGCACTTAACGCTTCAC	986
Oy	1138	aatgcgatgtatgtgcctgcctcccaagaatcgttccatctgtctatgtatgatatgacctc	1197
Db	987	AATGAGCTGATGGTGAACCTGTGTTCAAGGACCGCTTCATTTGCTGTGGTGGACATAGGCTTCT	1046
Oy	1198	ccaactacatcaccctccgaggggtgcctgctgcagacaacgaagctgtctgccaatgttga	1257
Db	1047	GGGACGACATCACTTTAGCGCGTGTCTGTGGTGGCCACCGGGCTGGCGTCATGTATGTA	1106
Oy	1258	gaacttgaatgaagaatcatgttcttcgatcatctgggagtagaactataaagiatggaa	1317
Db	1107	GACTTTTACACACAAGACATCGTGTCTGCTGCTGCTGCTGTCACAGACATCAAAAGCTGTGAC	1166
Oy	1318	acaagtaactgtgaaatttgtaaggaacctttaaataatgacacaagcagatcgctgtttg	1377
Db	1167	ACGAGCCACTGTAAATTTGTTGCTACTCTCAATGGGCGCAAGCGGGCATTTGCTGTCTC	1226
Oy	1378	cagtaacagggacacagcgtgtagtgaagtggtctcatctgcacaacactatcagaataggac	1437
Db	1227	CAGTACAGGAGATCGCTGTGTGTATGTGATATCAGATTAATFACCATTAAGGCTCTGGAT	1286
Oy	1438	atagaatgtgtgcatagttaacagtggttaagaagccatagagattgggtgctgtatt	1497
Db	1287	ATTGAATGTGTGCTCCTGTTTAAGAGTCCCTAGAGGACATGTGAAGATTTGTCACATGATC	1346
Oy	1498	cgatttataacaagaagatagtcagttggggcctctatgtgtaaaaataaagatgttggat	1557
Db	1347	CGGTTTGAATAAACAAGAGATTTGTCAGTGGGCTATGATGCGAAATTAAGTTGGGAC	1406
Oy	1558	cttgtgtgctgctttgacccccgtgtcctctgaagggaaacatctgtctcaagacccttgg	1617
Db	1407	TTTGCAACTCTCTTATACCTCGAGCCCGACGAACACATTTGTTTGGCCACATTTGGTG	1466
Oy	1618	gagcatccggaagatlltcgaactacagtttgatcatccagatctgacatgattga	1677
Db	1467	GAAATTTCTGACAGTGTGTTTGGGCTCCAGTTTATGATGATTGAGATCAGACAGCTCC	1526
Oy	1678	catgaatgaacaatcctcatctgtggaattcctaataatgatacagctgcaccaagtgaacc	1737
Db	1527	CATGATGACACTTTTGTGATTTGGGATTTCTTAAATGTGCTCCACAGTCCGCAATGTAG	1586
Oy	1738	ccccgttccccctctcgaacataacacctaagatcttcgaataataaaccataacgtacc	1796
Db	1587	ACCCGTTCTCTCCTCGAAACATPACCTTACATCTCTGAATPACAGTCTGCACATTTCAC	1645
RESULT 13			
AB033280			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;			
Bakallia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			

REFERENCE 1 (sites)
 AUTHORS Kojke, J., Sagara, N., Kirikoshi, H., Takagi, A., Miwa, T., Hirai, M. and
 TITLE Molecular cloning and genomic structure of the betaTRCP2 gene on
 JOURNAL chromosome 5q35.1
 MEDLINE Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
 REFERENCE 2 (bases 1 to 2252)
 AUTHORS Kato, M.
 TITLE Direct Submission
 JOURNAL Submitted (05-OCT-1999) Masaru Kato, National Cancer Center,
 Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
 (E-mail: m.kato@ncc.go.jp, Tel: 81-3-3542-2511 (ex. 4402),
 Fax: 81-3-3541-2685)

FEATURES
 SOURCE location/Qualifiers
 1. 2252
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /map="5q35.1"
 /tissue_type="fetal lung"
 157..1746
 /gene="TRCP2"
 157..1746
 /gene="TRCP2"
 /codon_start=1
 /product="F-box and WD-repeats protein beta-TRCP2 isoform
 B"
 /protein_id="BA02330.1"
 /db_xref="GI:7209611"
 /translation="MEPDSVIEDKTIELMNTSVEDONEDSPKKNLWISNCTSSV
 ILSRRPSEGNQKEDKICIFPDWSESDVEVEHLISMCHYQHINSLKPL
 ORDFTALPEGLDIAENILSYDARSLCAELVCKEMQVISEGLMKLIERMV
 TDPLMKGSEKRGMDQYLFKNRPDPGPSYRSLYPIIDIEIENMRCGHNIQ
 RIQCSSEKSKVCYCLQYDDEKILISLRNLSIKIMDKTSECLKVLGHTGSLQYD
 ERVYTGSSDSVVRWVNTGSEVNLTIHNEAVLHLEFSEGLWATGSKDSIAWMD
 ASADITLRVAVGRRAVNVDFPDKTIIVASGDRITKWNSTGCEVRVILNGKRC
 IACLOYRDLVAGSSDNTIRLMDIEGACILVEGHEILVRCILFMDKRIIVSAGID
 KIKVMDLQALDPRAPASTLCRLTLEHSGRFLQDFEPIIISSHDTLLIMDFLN
 VPSAONTNRSPTRYTYSR"

BASE COUNT 598 a 501 c 558 g 595 t

ORIGIN

Query Match 41.1%; Score 883.4; DB 9; Length 2252;
 Best Local Similarity 74.8%; Pred. No. 4.5e-243;
 Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

Db 631 GTACGACCTGATCCCTATGGAAGACTTTCAGAAAGAGGCTGGATACGACCTG 690
 Qy 718 ttcaaaacaacccctctgacggaaatgctccctcccaactctttatagaagcaattat 777
 Db 691 TTTAAAAACAGACCACCAAGATGCG-----CCTCAAAATTCATTTATAGCATATATAC 744
 Qy 778 cctaaattatagaagacattgagacataagaacttaattgagatgagaaagacatg 837
 Db 745 CCAAGATTTATCCAGGATATAGAGACTATGAAATTCATACCTGGGCTGTGGACACCAAC 804
 Qy 838 ttcaaggaatttaccgacgaagtgaagaaacgaaggaatttaccgtttacagatgat 897
 Db 805 TTGAGAGGATTCAGTCCCTCTGAAAATAGTAAAGTGTCTACTGTTTACAGTACGAT 864
 Qy 898 gatcaagaataagtaagcgccttcgagacaacaacatacaagatcctggatlaaacaca 957
 Db 865 GATGAAAAAATTTATCATGGCGCTACAGATTAATTCATTAAGATATGGGATTAACACAC 924
 Qy 958 ttggaatgcaagcgaattctcaagcgcatacagtgctcagtcctcgtctccagatgat 1017
 Db 925 CTGGAATGTTTGAAGGTGTTAACAGACACAGAGCTCTGCTCTGTCTGCAATGAT 984
 Qy 1018 gaagagatgatacataagacatcatcgattccacggtcgaagatgagatgagatgagat 1077
 Db 985 GAGGCTGATTTATGACTGCTCTTCAGATTTTACGCTGAGAGTGGATGATGAGAACAG 1044
 Qy 1078 ggtgaatgctaaacacggttgatcaccatgtgaaagcattctgcaacttggttcaat 1137
 Db 1045 GGTGAAGTCTTAAACATGATGATCCACACAAATGAGCTGTATGCACTTACGCTTACG 1104
 Qy 1138 aatgcatgatagtgacactgctccaaagatgcttccatctctgataagatgagatgag 1197
 Db 1105 AATGACATGATGAGCTGTTTCCAAAGACCGCTCCATTGCTGTGTGGACATGCGCTTC 1164
 Qy 1198 ccaactgacatcaccctccggaaggtgctggtcggacacgagctgctgtaaatgtgta 1257
 Db 1165 GCGACCGACATCATTATACGCGCGTCTCTGTTGAGCCACCGCGCTGCGCAATGATGTA 1224
 Qy 1258 gaacttgatgacaagtatactgttcttgcaactcgtggatgagatgagatgagatgag 1317
 Db 1225 GACTTTGACGACAACTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284
 Qy 1318 acaagatcttgtaatttgtaaggaacttaattgaacaaacgaagcattgctgctt 1377
 Db 1285 ACGACACCTGTGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1344
 Qy 1378 cagtaacaggaagcctggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1437
 Db 1345 CAGTACAGGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1404
 Qy 1438 atagaatgtgtgacatgcttcaagagtggttagaagccatggaattggtgctgtaatt 1497
 Db 1405 AATTGATGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1464
 Qy 1498 cgatttgatacaagaagataagtcagtggtggtggtggtggtggtggtggtggtggtggt 1557
 Db 1465 CGCTTTGATTAACAAGAGATGTGAGTGGGCTGATGATGGAATAATTAAGTTGGGAC 1524
 Qy 1558 cctgtgctgcttggaaccccggtgctcctgaggaagacactgctgctacggaacctgtg 1617
 Db 1525 TTGCAAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584
 Qy 1618 gaggacttcggaagaggttctgacatagatggttgaatccatccatgctcagtaattca 1677
 Db 1585 GAACATTTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1644
 Qy 1678 catgatacaacatacctcactcgtggaactcctaattgacgaagctgccaagctgaaccc 1737
 Db 1645 CATGATGACACTATTTGATTTGGATTTCTTAATGTGCTCCCAAGTCCCAAGATGAG 1704
 Qy 1738 cccggtccctctcgaacatacaccatacctcagtaataacatacatacagcagc 1796

```

Db      1705 ACCGCTTCCTCCGACAGAACATACACTTACATCTAGATACAGCTGCTGACTTTCACC 1763

RESULT  14
LOCUS   AB033281                2274 bp    mRNA    linear    PRI 14-MAR-2000
DEFINITION Homo sapiens BTRCP2 mRNA for F-box and WD-repeats protein isoform C, complete cds.
ACCESSION AB033281
VERSION   AB033281.1 GI:7209812
KEYWORDS  BTRCP2; F-box and WD-repeats protein beta-TRPC2 isoform C.
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
Molecular cloning and genomic structure of the betaTRCP2 gene on chromosome 5q35.1
Biochem. Biophys. Res. Commun. 269 (1), 103-109 (2000)
2 (bases 1 to 2274)
Katoch, M.
Direct Submission
Submitted (05-OCT-1999) Masaru Katoch, National Cancer Center,
Genetics Division, Tsukiji 5-chome, Chuo-ku, Tokyo 104-0045, Japan
(E-mail: mkato@nc.c.90.jp, Tel: 81-3-3542-2511 (ex. 4402),
Fax: 81-3-3541-2685)
FEATURES
         source          1..2274
             location/qualifiers
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /chromosome="5"
             /map="5q35.1"
             /issue_type="fetal lung"
             /gene="BTRCP2"
             /gene="BTRCP2"
             /codon_start=1
             /product="F-box and WD-repeats protein beta-TRPC2 isoform C"
             /protein_id="BA092331.1"
             /db_xref="GI:7209812"
             /translation="MEPDSVIEDKTELKSVPRSLMGLCANLVESMCALSCLOSMP
             VRCLQISNGTSVIVSRKRPSENGYOKEDLCIKYFDOMSEDOVEFEHLISRMCHY
             QHGHINSYLKPMLODFITALPQGLDHAENTLSTYLDASICAAEIVCKMORVISE
             GMLMKLIERMVRDPLMKGLSRRCMDQYLRNRPDPSPNSTRSLYKTIIDIEY
             IESNMRCGRHNRIOICRSENSKGYCLOIDDKLISGLRDSIKIWDTSLECLKLV
             TGHRTSVLCLODERIVITGSSDSVTRVWDVNTGEVLNLIHNEAVLHLRISNGLM
             TCSKDRSLAVMDASATDITLRVLYGHRAAVNVDFDKYIVSASGDDTIVMST
             CEFVRLNGHKRMAIACLOYRDLRVSSSDNTIRLMDIECGACRLVLEHEHLVNCIR
             FDKRIIVSAGYDKIKVMDLOALADPRASSTLCITLVEHSGRNFRLQDFEFOIIS
             SHDDTLIMDLVLPSPSAQNETRSPRTTYISR"

BASE COUNT      587 a      514 c      566 g      607 t

ORIGIN
Query Match      41.1%; Score 883.4; DB 9; Length 2274;
Best Local Similarity 74.8%; Pred. No. 4.5e-243;
Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1:

Db      413 GTGGAACATCTTATTTCACGAATGTGTCAATTATACGATGACATATTATACCTTACCG 472
Qy      478 aaacctatgttcagagagatttcatactacgtcttcgcagctcgggattgatacatc 537
Db      473 AAGCCCATGTTGCACGGCGGACTTTATTACCGCTTTACACAGAGAGCGCTTAGATTCACAA 532
Qy      538 gctgagacatctctgtataccttgatgacatgacatgacatgacatgacatgacatgac 597
Db      533 GCAGAAAAACATTTCTTTGTGACCTGAGATGCCAGTCTCTGTCTGCACAGACAGCTGGATGT 592
Qy      598 aaggaatggtacagagagagctctgtatgacatgacatgacatgacatgacatgacatgac 657
Db      593 AAAAAATGCGACGAGATGATCTAGAAAGAAATCTTTGGAAGAGCTGATTGAAACGAAG 652
Qy      658 gtcaagacagatctctgttgcagagagcctgagagagcctgagagagcctgagagagcct 717
Db      653 GTAGCGACATGATCCCGCTATGGAAGAGCTTTCCAGAAAGAAAGAGGATGATGATGATG 712
Qy      718 ttcaaaaacaaacctctcgtcagaggaatgctctcccaactcttltatagacacttatt 777
Db      713 TTTAAAAACAGACCACAGATGCG-----CCTCCAAATTTCATTTTATAGTCATTATAC 766
Qy      778 cctaaattatatacagacattgagacatagaaatcttaattgagatgagatgagatgagat 837
Db      767 CAAAAGTTTATCCAGATATAGAGACTATATGATTTAACTGGGGTGTGACGACACAAAC 826
Qy      838 ttcaagagaattcaactcgcgaatgtaaaacaaagaaagaaatgtaatttactgtttacagat 897
Db      827 TTCCAGAGATTCAGTCCGCTGTGAAATAGTAAGAGGTCTACGTATTACAGTACGAT 886
Qy      898 gatcagaagaatgtaagcgccttgagagacacacacacacacacacacacacacacacacac 957
Db      887 GATGAAAAAATTTATCATGTCGCTACAGATTAATTTATTAAATATGAGATTAACCAACG 946
Qy      958 ttggaatgcaagcgaatttcacagagcacaacacacacacacacacacacacacacacacac 1017
Db      947 CTGGAATGTTTAAAGTGTTAACAGTACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1006
Qy      1018 gagagagtgatcaatacagagatcaccgttccacagcagcagcagcagcagcagcagcagcag 1077
Db      1007 GACGCGTATGATTAACCTGCTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCAAG 1066
Qy      1078 ggtgaatgcaacacacacacacacacacacacacacacacacacacacacacacacacacac 1137
Db      1067 GGTGAAGTTCTTAAACATATGATTCACACACATGAGCGCTGATTTACACTTACGCTTACG 1126
Qy      1138 aatgcatgatagtgacacacacacacacacacacacacacacacacacacacacacacacac 1197
Db      1127 AATGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
Qy      1198 ccaactgacatcactcctcgcgaaggtgctgctgctgctgctgctgctgctgctgctgctgct 1257
Db      1187 GCGACCGACATACCTTACGCGGCTCTCTGTTGCGCACCGCGGCTCGCTGAAATGTGTA 1246
Qy      1258 gactttgataagatacattgttctgcatcgtgagataagataagataagataagataagata 1317
Db      1247 GACTTTGACGACAAATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1306
Qy      1318 acaagtaacttgtaatttgtaaggaacttaatgtaaggaacttaatgtaaggaacttgctgt 1377
Db      1307 ACGACACCTCTGTAATTTGTTCTGATCTCAATGCGCACCAAGCGGGGCTGCTGCTGCTG 1366
Qy      1378 cagtaacaggaacagcctgtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1437
Db      1367 CAGTACAGAGATGCGCTGCTGTTAGTGATGATGATGATGATGATGATGATGATGATGATGAT 1426
Qy      1438 atgaatgtagtgcagtgtaagagtgtaagagtgtaagagtgtaagagtgtaagagtgtaag 1497
Db      1427 ATTGAATGCTGCTGCTGCTGTTAAGAGTCTTAGAGGACATGAGAAATGTGCTGATGATC 1486
Qy      1498 cgattgataacagaagagatagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 1557
Db      1487 CGGTTTGAATTAACAGAGATTTGCTAGTGGGGCTTATGATGAGAAATTAAGTTTGGGAC 1546

```


QY 1558 ctgtgtgctgttctgaccccgctcctcgaaggagacactctgtctacagaccctgtg 1617
 Db 1547 TTGCAAGCTCTCTTACCTTCAGCCCGCCAGCAACATTTGGTTTGGCCATTTGGTG 1606
 QY 1618 gacattccgagagatttctgacagtttgatgaatccagattgtcgaatgtca 1677
 Db 1607 GAACATTTCTGGACGCTGTTCGGCTCCAGTTGATGATTCATCATCAGCAGCTCC 1666
 QY 1678 catgatgacaacatccatctcgtggaactcctaataatgatacagctgcccagaatgaacc 1737
 Db 1667 CATGATGACACTATTGATTTGGATTTCTTAATGTGCTCCCACTGCCCGCAATGAG 1726
 QY 1738 cccgttccccctctcgaac 1796
 Db 1727 ACCCGTTTCCCTCCAGCAACATACACTTACATCTAGATTAACAGTCTGCACCTTCACC 1785

RESULT 15

AB014596 4230 bp mRNA linear PRI 06-FEB-1999
 LOCUS Homo sapiens mRNA for KIAA0696 protein, partial cds.
 DEFINITION
 ACCESSION AB014596
 VERSION
 KEYWORDS
 SOURCE Homo sapiens adult male brain cDNA to mRNA, clone.lib:pbluescriptII SK plus clone:HK04053.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 4230)
 Ohara, O., Suyama, M., Nagase, T. and Ishikawa, K.
 Direct Submission
 Submitted (26-MAY-1998) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, Yana 1532-3, Kisarazu, Chiba
 292-0812, Japan (E-mail:cdna@fokazusa.or.jp, Tel:+81-438-52-3913,
 Fax:+81-438-52-3914)
 2 (sites)
 Ishikawa, K., Nagase, T., Suyama, M., Miyajima, N., Tanaka, A.,
 Kotani, H., Nomura, N. and Ohara, O.
 Prediction of the coding sequences of unidentified human genes. X.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro
 JOURNAL Mammalian Genome 9 (3), 169-176 (1998)
 MEDLINE 98403880
 FEATURES

REFERENCE

AUTHORS

TITLE

Location/Qualifiers
 1. 4230
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HK04053"
 /sex="male"
 /tissue="brain"
 /clone_lib="pbluescriptII SK plus"
 /dev_stage="adult"
 1. 1654
 /gene="KIAA0696"
 <1. 1654
 /gene="KIAA0696"
 /codon_start=2
 /product="KIAA0696 protein"
 /protein_id="BAA31671.1"
 /translation="RGGAAGAAMPDSVIEDKTELMCSVPSRLIMGICANVESMCAL
 SCLOMPSVRLIOISNGTSVIVSRKRPSEGNOKERDICIKYDONSQSEVFEVH
 LISRMGHQGHINSYLAKPMQGRPTALPEQGLDHAENLSTLDARSCLAEVLCK
 EMORVISEGMKMLITERVITDPLMGSLSRKWDYLLFNRPIDGPNSEFVSLVP
 KIITIDETIESNMRGRHNLQROCKRENSGVYCLDYDEKITSGLRNSIKIMDT
 SLLEKLVHTGTSVLCLOYDERVIVTSSSDSVRWVDVNTGEVLTILHNEVNLHL
 RFSNGLAVTCSKDRSLAVMDMASATDITLRVLVGHRAAVVDFDKYIYSAGDRT
 IKWVSTCEVTRTNGHKGRLACIOLRDRLVYSSGSDNTLRIMDIEGALRTIECH
 EELVRCIRPNKRIVSGAYDKIKVMPLQALIDPRASITCLTFLVEHSGRVRLQF
 DEPOLISSHDDTLIMDFLAVPPSAONETRSRSTYTVISR"

source

CDS

gene

BASE COUNT

1121 a 848 c 965 g 1296 t

ORIGIN

Query Match 41.1%; Score 883.4; DB 9; Length 4230;
 Best Local Similarity 74.8%; Pred. No. 4,9e-243;
 Matches 1122; Conservative 0; Mismatches 371; Indels 6; Gaps 1;

QY 298 aatggcactccagatgattgtgcccgaagcaagcaactcctcagaagctatgaag 357
 Db 179 AATGGAACATCATCTGTGATCGTCTCCAGAAAGAGGCCATCAGAAAGAACTATCAAAA 238
 QY 358 gaaagaagactgtgtgtcaaatcatttgagcagtggtgtcagaatgaatgaat 417
 Db 239 GAAAGAGACTTGTGATTAATTAATTTGACAGAGGCTGTAATCATGATCAAGGAAATTT 298
 QY 418 gttgacatcttatcccaaatgtgtcatattaccacaatgtggacaataactcgatc 477
 Db 299 GTGGAACTCTTATTTTCAGCAATGTGTCAATTAACATGACATATTAATTAATTAAT 358
 QY 478 aaacctatgtcagaagatttcataactgtctcgaactcgtgagatgtgatatc 537
 Db 359 AAGCCCATGTTGACGCGGAGCTTATTACCGCTTTACCAAGCAAGGCTTATGATCATATA 418
 QY 538 gctgagaaactctgttcatatcctgtgataaactcaatctgtgtgtgtgtgtgtgtgt 597
 Db 419 GCAGAAACATTTCTTTCGATCCGATGACGAGGCTCTGTGTGACGAGAGCTGTATGT 478
 QY 598 aaggaatgtgacagatgtgacactgtgatagtgtgtgtgtgtgtgtgtgtgtgtgtgt 657
 Db 479 AAGGATGCGAGGAGATGATCTAGAGAGGATCTTTGGAAGAGCTGATGAGCAATG 538
 QY 658 gtcaagacagatctctgtgtgagagccgtgtgcaagcaagagatgtgtgtgtgtgtgt 717
 Db 539 GTACGACATGATCCCTATGGAAGAGCTTTGAGAAAGAGGAGGAGTGTGTGTGTGTGT 598
 QY 718 ttaaaaaaacaactctcctcaggaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 777
 Db 599 TTTTAAACAGACCCACAGATGCG-----CCTCCAAATTCATTTATAGTATTATATAC 652
 QY 778 cctaaattatacaagacattgagacatgaatcctcaatgtgtgtgtgtgtgtgtgtgtgt 837
 Db 653 CCNAAGATTATCCAGATATAGACTATAGATTAATGATTAATGATGATGATGATGATGAT 712
 QY 838 ttcaagagatctcctgt 897
 Db 713 TTGCAGAGATTCAGTGTGCGCTCTGAAATAGTAAAGGTGTGTGTGTGTGTGTGTGTGT 772
 QY 898 gatcagaataatgtaagcgtctgtgagcaacacacacacacacacacacacacacacac 957
 Db 773 GATGAAATTAATATCATGTGCTACAGAGATTAATTTAATTAATTAATTAATTAATTAAT 832
 QY 958 ttggaatgcaagcgaattctcacaagcctacaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1017
 Db 833 CTGGAATGTTGAAAGTGTAAACAGACACACAGCGCTGTCTGTCTGTCTGTCTGTCTGT 892
 QY 1018 gaggagatgatacaacaagatcattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1077
 Db 893 GAGCGGTGATTTGATCTGCTCTTCAATTTACAGTGTAGAGTGTGGATGTGAACAGC 952
 QY 1078 gttgaatgtcacaacgctgt 1137
 Db 953 GGTGAAGTCTTAAACATGATTAATCAACCAATGAGCGCTGTGTGTGTGTGTGTGTGTGT 1012
 QY 1138 aatgcatatgt 1197
 Db 1013 AATGACATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1072
 QY 1198 ccaactgtcatcactcctcgaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1257
 Db 1073 GTGACCGATATCTTACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1132
 QY 1258 gacttctgataaagataatgt 1317

Sat May 11 12:50:53 2002

us-09-601-168a-1.rge

Page 20

Db	1133	GACCTTCACGACAAAGTACACGTCGTCTGGCTCTGGTGGACAGGACCATCAAAAGTCGAGAC	1192
Qy	1318	acaagtaacttgtaatttgtaaggaccttaatlgacacaaagagcattgcctgtttg	1377
Db	1193	ACGAGCCACCTGTAATTTGTTCCCTACTCTCAATGGGCAACACGGCGATTGCTGTCTC	1252
Qy	1378	cagtaacagggacagcgctcgtagtgagtgctcatcagcaaacacatcagataatgagc	1437
Db	1253	CAGTACAGGGATCGCCGTGTGTATGGATCTCAGATAAATACATTAGCTCTGGGAT	1312
Qy	1438	atagaatgtagtcatgttttaagagtgttagaagccatlgaggaattgtagctgtatt	1497
Db	1313	ATTGAATGTGTGCTGCTGTTTAAAGTCCTAAGAGGGACATGAAGAAATGGTCGATGCATC	1372
Qy	1498	cgatttgataacaaagagatagtcacatggggccatagatggaanaataaagtgtagat	1557
Db	1373	CGGTTTATATACAGAGGAGATTGTCACTGGGCTTATGATGGGAAAATTAAATTGGGAC	1432
Qy	1558	cttgtagctgactltgagaccccgtagtctctgcagagagacactgltctagagcccttg	1617
Db	1433	TTGCAACGTGTGCTTTGACCTCCAGCCCCGACAGACAAATGTGTTGGCCCATTTGGTG	1492
Qy	1618	gagcattccgggaagatttttgactacagtttgatgaattccagatttgtaagtaagtcca	1677
Db	1493	GAACATTCTGGACGTGTGTTTGGGCTCCGATTTGATGACTTTTCAATCTGCAAGCTCC	1552
Qy	1678	catgatgacaacatccctcatcttggaacttccatagatgcacagctgcccagaatgacc	1737
Db	1553	CATGATGACACTATTTTGATTGTGGATTCTTAAATGTGCTCCGCCAGAAATGAG	1612
Qy	1738	ccccgttcccctctctggaacatacactataatctcaagataaataaccaataacgagc	1796
Db	1613	ACCGTTCTCTCCCTCCAAACAAATACACTTCAATCTTAGAATACAGTCTGCACTTTCACC	1671

Search completed: May 11, 2002, 09:15:52
Job time: 14986 sec

PR 09-DEC-1998: 98FR-0015545.
 PR 30-JAN-1998: 98FR-0001100.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 XX Arenzana Seisdedos F, Benarous R, Concorde J, Durand H;
 PI Kroll M, Margolin F,
 XX
 DR MPI, 1999-469329/39.
 DR P-PSDB; AA124054.
 XX
 PT New human beta-transducin repeat containing protein and its
 fragments useful as, or to screen for, antiviral, antitumour,
 anti-inflammatory and anti-Alzheimer's agents
 XX
 PS Claim 7: Page 57-60; 71pp: French.
 XX
 CC The present sequence encodes a human beta-transducin repeat containing
 CC protein (beta-Trcp). The protein directs proteins to the proteosome
 CC degradation pathways. The protein is able to interact with the Ypu
 CC protein of human immune deficiency virus-1 (HIV-1), cellular proteins
 CC IkappaB or beta-catenin (bc) and/or protein Skp1. The protein controls
 CC ubiquitinylation of phosphorylated proteins and thus their targeting to
 CC proteosomes for degradation. Depending on whether the process is
 CC inhibited or promoted, the result may be delayed breakdown of CD4 (in
 CC cases of HIV-1 infection); increased activity of IKB (and thus reduced
 CC activity of NFkappaB) and increased degradation of mutant bc in tumour
 CC cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
 CC patients. The beta-Trcp protein, and its active peptide fragments, or its
 CC nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
 CC antitumour agents that disrupt cell cycle regulation or protein
 CC degradation in human tumour cells, and anti-inflammatory agents that
 CC disrupt activation by NFkappaB. Fragments of the protein are also
 CC useful for treating osteo-articular inflammation or acute inflammation
 CC associated with release of tumour necrosis factor.
 CC
 XX
 SQ Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other:

Query Match 100.0%; Score 2151; DB 20; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgctgacacaaaggcgcccgcgagagcgagaccagtgcc 60
 DB 1 tgcgttgctgcgctgacacaaaggcgcccgcgagagcgagaccagtgcc 60
 QY 61 tgcgcatatggaaccgcgcgagcggtgctgcaagagaagacccaattatgaat 120
 DB 61 tgcgcatatggaaccgcgcgagcggtgctgcaagagaagacccaattatgaat 120
 QY 121 tccctcagagagaagactgttaataatggcgaaccctcagagaataatacagaag 180
 DB 121 tccctcagagagaagactgttaataatggcgaaccctcagagaataatacagaag 180
 QY 181 aattcactaagcagaacatacaacagctgtgccaagctctttaaaccagaacagta 240
 DB 181 aattcactaagcagaacatacaacagctgtgccaagctctttaaaccagaacagta 240
 QY 241 tatttagcaagcactgtcctaagaactgtgtgccccaaacaaacttccaat 300
 DB 241 tatttagcaagcactgtcctaagaactgtgtgccccaaacaaacttccaat 300
 QY 301 ggcactccagatgatgtgtgcccagaacggaactctcagcaagctatgaagaagaa 360
 DB 301 ggcactccagatgatgtgtgcccagaacggaactctcagcaagctatgaagaagaa 360
 QY 361 aaggaacgtgtgtaataactttgagcagtggtcagaagtcagaatcaagtggaattgtg 420
 DB 361 aaggaacgtgtgtaataactttgagcagtggtcagaagtcagaatcaagtggaattgtg 420
 QY 421 gaaactctatatacccaaatgtgtcattacaacatgagcacaataaactgtatcttaaa 480

DB 421 gaaactctatatacccaaatgtgtcattacaacatgagcacaataaactgtatcttaaa 480
 QY 481 cctatgttcagagagatcttcaactgctgtccagctcggggattggatcatcgct 540
 DB 481 cctatgttcagagagatcttcaactgctgtccagctcggggattggatcatcgct 540
 QY 541 gagaacatctgtcactacgtgatgcgaataactactatgtgtcgtcgaactgtgtgcaag 600
 DB 541 gagaacatctgtcactacgtgatgcgaataactactatgtgtcgtcgaactgtgtgcaag 600
 QY 601 gaaatgttcagagagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 660
 DB 601 gaaatgttcagagagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 660
 QY 661 aggaacagatctcctgtgtgagagcgctgcaagaagagatgtgagacagttatttc 720
 DB 661 aggaacagatctcctgtgtgagagcgctgcaagaagagatgtgagacagttatttc 720
 QY 721 aaaaacaaactcctgacgggaatgtccctcccaactctttatagaacattatcct 780
 DB 721 aaaaacaaactcctgacgggaatgtccctcccaactctttatagaacattatcct 780
 QY 781 aaaaattacagaagcattgagacacatagaaatctaatgtgagatgtggaacatagtta 840
 DB 781 aaaaattacagaagcattgagacacatagaaatctaatgtgagatgtggaacatagtta 840
 QY 841 cagagaatctacgtgcggaagtgtgaacaagaagaagttacgtttacagttatgat 900
 DB 841 cagagaatctacgtgcggaagtgtgaacaagaagaagttacgtttacagttatgat 900
 QY 901 cagagaatctacgtgcggaagtgtgaacaagaagaagttacgtttacagttatgat 960
 DB 901 cagagaatctacgtgcggaagtgtgaacaagaagaagttacgtttacagttatgat 960
 QY 961 gaaatgttcagagagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1020
 DB 961 gaaatgttcagagagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1020
 QY 1021 aagatgttcatacagaagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1080
 DB 1021 aagatgttcatacagaagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1080
 QY 1081 gaaatgttcatacagaagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1140
 DB 1081 gaaatgttcatacagaagatctcctgatgcatgtgtggaagaagcttaccggaagaatgtgc 1140
 QY 1141 ggcattgtgtgacgtgtcccaagaatcgttccatgtgtgtatgtatgtgcctccca 1200
 DB 1141 ggcattgtgtgacgtgtcccaagaatcgttccatgtgtgtatgtatgtgcctccca 1200
 QY 1201 actacatcaccctccggaaggtgtgtgtggaacacagctcgtcgaatgtgtgagac 1260
 DB 1201 actacatcaccctccggaaggtgtgtgtggaacacagctcgtcgaatgtgtgagac 1260
 QY 1261 ttgtatgacaagaatcattgttctgtcatctggtggaatagaactaagaagtatggaacaca 1320
 DB 1261 ttgtatgacaagaatcattgttctgtcatctggtggaatagaactaagaagtatggaacaca 1320
 QY 1321 agtaactgtgaatcttgaagacctaataatggcacaagaagagcattgctgtttgacg 1380
 DB 1321 agtaactgtgaatcttgaagacctaataatggcacaagaagagcattgctgtttgacg 1380
 QY 1381 taacaggaacagcgtgtgtgagtggtcctacgcgaacacatacagaatatagggaacata 1440
 DB 1381 taacaggaacagcgtgtgtgagtggtcctacgcgaacacatacagaatatagggaacata 1440
 QY 1441 gaatgtgtgtcagttacagagtggttagaagcagatggaatgtgtgctgtgtatcga 1500
 DB 1441 gaatgtgtgtcagttacagagtggttagaagcagatggaatgtgtgctgtgtatcga 1500
 QY 1501 ttgtatgaacagaagagatgtcagttggtggtcctatgtatggaaatgaatgtgtggtatcct 1560

QY 541 gagaacatctgtcacaacctgagatgcacaatacctaagtctgtcgtgaacttctgtgcaag 600
 |||||
 Db 541 gagaacatctgtcacaacctgagatgcacaatacctaagtctgtcgtgaacttctgtgcaag 600
 QY 601 gaatgtacccagatgaaacctctgaatgcatgctgtggaagaagcttaccagaaatgctc 660
 |||||
 Db 601 gaatgtacccagatgaaacctctgaatgcatgctgtggaagaagcttaccagaaatgctc 660
 QY 661 agaacgattctctgtggaagaagcctgcgagaaagaaagaaatggaagaaatattatc 720
 |||||
 Db 661 agaacgattctctgtggaagaagcctgcgagaaagaaagaaatggaagaaatattatc 720
 QY 721 aaaaacaacctctgacgggaatgtctctcccaactctttatagagacattatcct 780
 |||||
 Db 721 aaaaacaacctctgacgggaatgtctctcccaactctttatagagacattatcct 780
 QY 781 aaaaattatacaagaacttgaagaataagaatactaatltggaagatgtggaagacattta 840
 |||||
 Db 781 aaaaattatacaagaacttgaagaataagaatactaatltggaagatgtggaagacattta 840
 QY 841 cagaagattcaactgcgcgaagtgaaacaagcaaaaggaatgttactgtttacagatagat 900
 |||||
 Db 841 cagaagattcaactgcgcgaagtgaaacaagcaaaaggaatgttactgtttacagatagat 900
 QY 901 cagaanaatgtaagcgccctcgagacaacacaatcaaatcaagatctggataaacaacatg 960
 |||||
 Db 901 cagaanaatgtaagcgccctcgagacaacacaatcaaatcaagatctggataaacaacatg 960
 QY 961 gaatgaagcgaattctgaagaagcctaagaattcagttcctctgtctccagatagatgag 1020
 |||||
 Db 961 gaatgaagcgaattctgaagaagcctaagaattcagttcctctgtctccagatagatgag 1020
 QY 1021 agagtgatacaacaagatcagatctggaatccacggtcagaagtgaggatgataaatacagt 1080
 |||||
 Db 1021 agagtgatacaacaagatcagatctggaatccacggtcagaagtgaggatgataaatacagt 1080
 QY 1081 gaaatgctaaacaagcttgattacacatttgaagaagcagctctcgaacttgcgtttcaata 1140
 |||||
 Db 1081 gaaatgctaaacaagcttgattacacatttgaagaagcagctctcgaacttgcgtttcaata 1140
 QY 1141 ggcatagtatgtgacctgcgcacaagaatcgttccatgtcgtatggaatagtcctcccca 1200
 |||||
 Db 1141 ggcatagtatgtgacctgcgcacaagaatcgttccatgtcgtatggaatagtcctcccca 1200
 QY 1201 actgacataacctcccgagagtgctgtgcgagacacagcgtcgtcaatgtttagac 1260
 |||||
 Db 1201 actgacataacctcccgagagtgctgtgcgagacacagcgtcgtcaatgtttagac 1260
 QY 1261 ttgtagacaagaatgatttcttctcaatctgagagtagaactataaagtagaagaaca 1320
 |||||
 Db 1261 ttgtagacaagaatgatttcttctcaatctgagagtagaactataaagtagaagaaca 1320
 QY 1321 agtactgtgaaattgtgaagaccttaaatgagacaacaacagagcattgcctgttgcag 1380
 |||||
 Db 1321 agtactgtgaaattgtgaagaccttaaatgagacaacaacagagcattgcctgttgcag 1380
 QY 1381 tacaaaggacagcgtgtagtgagtgctcaatctgacaacacatcagattatgagacata 1440
 |||||
 Db 1381 tacaaaggacagcgtgtagtgagtgctcaatctgacaacacatcagattatgagacata 1440
 QY 1441 gaatgtgtgcaatgttttcagagtggttaagaagcctaaggaattggtgctgtattatcga 1500
 |||||
 Db 1441 gaatgtgtgcaatgttttcagagtggttaagaagcctaaggaattggtgctgtattatcga 1500
 QY 1501 ttgataacaagaagatagatcagtgaggcctatgataagtaaaatlaaagttgagatctt 1560
 |||||
 Db 1501 ttgataacaagaagatagatcagtgaggcctatgataagtaaaatlaaagttgagatctt 1560
 QY 1561 gtgagtgcttcttgagaccctcgctcctcgagaggaacactgttctacagacacctgtgag 1620
 |||||
 Db 1561 gtgagtgcttcttgagaccctcgctcctcgagaggaacactgttctacagacacctgtgag 1620
 QY 1621 catcccgagaagatgttttcgactacagtttgatgaattccagattgtcagtagttacaat 1680
 |||||

Db 1621 catcccgagaagatgttttcgactacagtttgatgaattccagattgtcagtagttacaat 1680
 |||||
 QY 1681 gatgacaacatcccatcattgtgagcttccataatgatccagctgcgcaagctggaaccccc 1740
 |||||
 Db 1681 gatgacaacatcccatcattgtgagcttccataatgatccagctgcgcaagctggaaccccc 1740
 QY 1741 cgttccctcttcgaaacatacactatcctcagataataaataacatacactgacctcat 1800
 |||||
 Db 1741 cgttccctcttcgaaacatacactatcctcagataataaataacatacactgacctcat 1800
 QY 1801 acttgcccaagaccatlaaagttgcggtatattagctatctgcgaatcaccagatgagc 1860
 |||||
 Db 1801 acttgcccaagaccatlaaagttgcggtatattagctatctgcgaatcaccagatgagc 1860
 QY 1861 aacacagtgaaacatacactacactgcccagtttccctgagctagccgagagcaggctt 1920
 |||||
 Db 1861 aacacagtgaaacatacactacactgcccagtttccctgagctagccgagagcaggctt 1920
 QY 1921 tgagactcctgttggagacaagttgtctcagctgcgagcccaagacggttctaccagaca 1980
 |||||
 Db 1921 tgagactcctgttggagacaagttgtctcagctgcgagcccaagacggttctaccagaca 1980
 QY 1981 actgactgctcagtgctgtcatcagaagatgtctctatcaatgttgaatgttgaac 2040
 |||||
 Db 1981 actgactgctcagtgctgtcatcagaagatgtctctatcaatgttgaatgttgaac 2040
 QY 2041 tttaaacctccctcctcctcctccttccacacctgcagctagtttttccattgttc 2100
 |||||
 Db 2041 tttaaacctccctcctcctcctccttccacacctgcagctagtttttccattgttc 2100
 QY 2101 cagacaagtgactataaataatattagttgttccagaaaaaaa 2151
 |||||
 Db 2101 cagacaagtgactataaataatattagttgttccagaaaaaaa 2151
 |||||

RESULT 3
 AA293350
 ID AA293350 standard; cDNA; 2151 BP.
 XX
 AC AA293350;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE Sequence encoding F-box protein FBP-1.
 XX
 KW F-box protein; FBP; diagnosis; treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 70..1779
 FT /*tag= a
 FT /product= FBP-1
 PN W0200012679-A1.
 XX
 PD 09-MAR-2000.
 XX
 PF 27-AUG-1999; 99WO-US19560.
 XX
 PR 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 99US-0118568.
 PR 15-MAR-1999; 99US-0124469.
 XX
 PA (UYNV) UNITV NEW YORK STATE.
 XX
 PI Chaur DS, Pagano M, Latres E;
 XX

Db	1741	cgctccccccttcggaataacacccctcacctccagaataataacataacacagactcat	18000
Qy	1801	acttgcgccagagaccattaaagtttgtagtattaaagtaactcgcacaataaccagagatgac	18600
Db	1801	acttgcgccagagaccattaaagtttgtagtattaaagtaactcgcacaataaccagagatgac	18600
Qy	1861	aacacacgataacatacaactactgcgccagtttcctctgactacgcgcagagagagagctt	19200
Db	1861	aacacacgataacatacaactactgcgccagtttcctctgactacgcgcagagagagagagctt	19200
Qy	1921	tgaagaccctcgtttgagacacagcttggctgcgcagctgcgcgcacagacagcgtactacagaca	19800
Db	1921	tgaagaccctcgtttgagacacagcttggctgcgcagctgcgcgcacagacagcgtactacagaca	19800
Qy	1981	actgacacgcttcagtgctcgtctatccagaagaatgtctctatacaaatggaatgatggaac	20400
Db	1981	actgacacgcttcagtgctcgtctatccagaagaatgtctctatacaaatggaatgatggaac	20400
Qy	2041	ttttaaaccccccctccctcctcccttcacacccgcagacacagatgttttcccatatggttc	21000
Db	2041	ttttaaaccccccctccctcctcccttcacacccgcagacacagatgttttcccatatggttc	21000
Qy	2101	cagacaaaggtgactataataataattgagtgttttgcgcagaaaaaaaaa	2151
Db	2101	cagacaaaggtgactataataataattgagtgttttgcgcagaaaaaaaaa	2151

CC 1)ases) which can be used for the targetted degradation of a target
CC polypeptide in vivo. Targetted degradation is achieved by expressing
CC the ubiquitin ligase in a cell linked to the interaction domain of
CC the target polypeptide and thereby recruiting the target polypeptide
CC to the ubiquitin ligase. Such methods are useful for decreasing or
CC increasing the level of a target polypeptide and for creating and
CC expressing a destabilized polypeptide which is subjected to SCF
CC mediated proteolysis. Degrading any desired protein in a cell is
CC useful for preventing or treating diseases caused by the presence of
CC abnormal amount of the specific polypeptides, for drug discovery and
CC for gene therapy. Diseases treated include cancer, by degradation of
CC oncoproteins, Huntington's disease, other proliferative disorders and
CC microbial infections. The method provides a quick and easy
CC alternative to gene knockout technology. The target polypeptide can
CC be degraded at all stages, or a specific stage, of development in the
CC mature animal.
XX
XX Sequence 2151 BP; 628 A; 467 C; 513 G; 543 T; 0 other;

Dh	721	aaaaaactctctgacgagaaatgctctctcccaactctttttataagagcaattaccct	760
Qy	781	aaaaatatacaaaacattgagacaatagaatctaaattgagagatgtagaagacatagtta	840
Dh	781	aaaaattacaagaacattgagacaatagaatctaaattgagagatgtagaagacatagtta	840
Qy	841	cagagaaattcaacgcccgaagtgaaacaagaacgaagagtttaactgttaccatgatgatg	900
Dh	841	cagagaaattcaacgcccgaagtgaaacaagaacgaagagtttaactgttaccatgatgatg	900
Qy	901	cagagaaatagtaagcggcctctcgagacaagaacacatcaagatctcggatataaacaacatg	960
Dh	901	cagagaaatagtaagcggcctctcgagacaagaacacatcaagatctcggatataaacaacatg	960
Qy	961	gaatgcgaagcgaaattctcaaaagccataaagtttaagtcctctgctccagtatgatgag	1020
Dh	961	gaatgcgaagcgaaattctcaaaagccataaagtttaagtcctctgctccagtatgatgag	1020
Qy	1021	agagtgatcatacaagagatcatcgcgatcttcacagcgtacagatgtagagatgataacagat	1080
Dh	1021	agagtgatcatacaagagatcatcgcgatcttcacagcgtacagatgtagagatgataacagat	1080
Qy	1081	gaaatgcataaacaacgcttgatctcaaccatgtgtgaagcaagctctgcactctgctttcaataat	1140
Dh	1081	gaaatgcataaacaacgcttgatctcaaccatgtgtgaagcaagctctgcactctgctttcaataat	1140
Qy	1141	ggcatgtagtgtagacctctccaaagaatcttcacattgctgatacttgagatataagccctccca	1200
Dh	1141	ggcatgtagtgtagacctctccaaagaatcttcacattgctgatacttgagatataagccctccca	1200
Qy	1201	actgacataccctcccgagagcgtcgtcgtctgcagacacggagcgtcgtccaaatgtgtagac	1260
Dh	1201	actgacataccctcccgagagcgtcgtcgtctgcagacacggagcgtcgtccaaatgtgtagac	1260
Qy	1261	tttgatgaagaagatgatactgtttcttgatattctggagatgaagaacatataaagatataagaaaca	1320
Dh	1261	tttgatgaagaagatgatactgtttcttgatattctggagatgaagaacatataaagatataagaaaca	1320
Qy	1321	agtaactgtgaaattgtgaagagaccttaaaatggaacacaaacggagcattgtccctgtttgag	1380
Dh	1321	agtaactgtgaaattgtgaagagaccttaaaatggaacacaaacggagcattgtccctgtttgag	1380
Qy	1381	tacagggacagagcctcgtagtagtggtctcaatctcgaaacaacatcataagattatgggacata	1440
Dh	1381	tacagggacagagcctcgtagtagtggtctcaatctcgaaacaacatcataagattatgggacata	1440
Qy	1441	gaatgtgtgatagttttaacagagtgcttaagaagcgcataaggaatctgtgtgtatgatacga	1500
Dh	1441	gaatgtgtgatagttttaacagagtgcttaagaagcgcataaggaatctgtgtgtatgatacga	1500
Qy	1501	tttgataacaagaagagatagtcagtgvgggccatgatgatgaaaaataaagtgtagactt	1560
Dh	1501	tttgataacaagaagagatagtcagtgvgggccatgatgatgaaaaataaagtgtagactt	1560
Qy	1561	gtgagctgttttgagcccccgtgcctccctgcaggggaacactgtgtctaaaggccctgttgag	1620
Dh	1561	gtgagctgttttgagcccccgtgcctccctgcaggggaacactgtgtctaaaggccctgttgag	1620
Qy	1621	catttcggaagaagtttttcgatactacagtttgatgaaattccacagatgtcagatgtltaacat	1680
Dh	1621	catttcggaagaagtttttcgatactacagtttgatgaaattccacagatgtcagatgtltaacat	1680
Qy	1681	gatgacgaacaatccctcaatctcagacttccctaataatgataccagctcggcccaagctgaaaccccc	1740
Dh	1681	gatgacgaacaatccctcaatctcagacttccctaataatgataccagctcggcccaagctgaaaccccc	1740
Qy	1741	gattccccccttcgtaacaatcacactaacatctccagaataaatacacaataacgagactcat	1800
Dh	1741	gattccccccttcgtaacaatcacactaacatctccagaataaatacacaataacgagactcat	1800
Qy	1801	acttgcgaaggaaccatlaaagtgtcggtatattaacgtatactgcgaataacgaagtagagc	1860

Query Match 100.0%; Score 2151; DB 22; Length 2151;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tgcgttgctgcgctgcgcacccaagggcgcccgcgagagcgagccagctggcc 60
 DB 1 tgcgttgctgcgctgcgcacccaagggcgcccgcgagagcgagccagctggcc 60

QY 61 tggcgattatggacccggcgagggggtgctgcgaagagagagcactgaattatgaat 120
 DB 61 tggcgattatggacccggcgagggggtgctgcgaagagagagcactgaattatgaat 120

QY 121 tccctcagagagagagagactgtaataatggcgaccccccaggaagataatacagagag 180
 DB 121 tccctcagagagagagagactgtaataatggcgaccccccaggaagataatacagagag 180

QY 181 aattcaacttagacagacatacaacagctgctgcagactctgcttaacccaagaacagta 240
 DB 181 aattcaacttagacagacatacaacagctgctgcagactctgcttaacccaagaacagta 240

QY 241 tcttttagcaagacactgctatgaagactgagaaattgtggtgccaacaacaacttgcacat 300
 DB 241 tcttttagcaagacactgctatgaagactgagaaattgtggtgccaacaacaacttgcacat 300

QY 301 ggcacttccagatgattgtgcccagacagcaaaactctcagcagactatgaaagaa 360
 DB 301 ggcacttccagatgattgtgcccagacagcaaaactctcagcagactatgaaagaa 360

QY 361 aagggaactgtgttaaaactttgagcagctgctcagagctcagatcaagtggaactgtg 420
 DB 361 aagggaactgtgttaaaactttgagcagctgctcagagctcagatcaagtggaactgtg 420

QY 421 gaacactataatacccaaatgtgcatatacaacaatggcacaataaactcgtatcttaa 480
 DB 421 gaacactataatacccaaatgtgcatatacaacaatggcacaataaactcgtatcttaa 480

QY 481 cctatgttgcagagagatttcaatactgctctgcagactcgggagttgatalcatactgct 540
 DB 481 cctatgttgcagagagatttcaatactgctctgcagactcgggagttgatalcatactgct 540

QY 541 gagaacattctgtactatgctgaatgccaatcaactgtgctgcgaactgtgtgcaag 600
 DB 541 gagaacattctgtactatgctgaatgccaatcaactgtgctgcgaactgtgtgcaag 600

QY 601 gaattgtaccgagtgacctctgatgtgcactgtctgtgaaagagcttatcgagagatgtc 660
 DB 601 gaattgtaccgagtgacctctgatgtgcactgtctgtgaaagagcttatcgagagatgtc 660

QY 661 aggcagagattctctgtgagagagcctgcgcagaaacgaagagatgaggaagatattatc 720
 DB 661 aggcagagattctctgtgagagagcctgcgcagaaacgaagagatgaggaagatattatc 720

QY 721 aaaaacaaacctcctgcagcggaatgctctcccaactcttttlaagagcacttatcct 780
 DB 721 aaaaacaaacctcctgcagcggaatgctctcccaactcttttlaagagcacttatcct 780

QY 781 aaaaattatacagagacttgagacaaatagatctaatitgagagatgctgaaagacatagttta 840
 DB 781 aaaaattatacagagacttgagacaaatagatctaatitgagagatgctgaaagacatagttta 840

QY 841 cagaagattcactgcgcagagatgaaacaaagcaagagagtttctctgttatacagatagat 900
 DB 841 cagaagattcactgcgcagagatgaaacaaagcaagagagtttctctgttatacagatagat 900

QY 901 cagaagattagtaagcgagcctcgcagacacacaactcaagactctggaataaacaacttg 960
 DB 901 cagaagattagtaagcgagcctcgcagacacacaactcaagactctggaataaacaacttg 960

QY 961 gaatgcaaggaattctcaacagccatacaggttcaagtctctctgtctccagatagatgag 1020
 DB 961 gaatgcaaggaattctcaacagccatacaggttcaagtctctctgtctccagatagatgag 1020

QY 1021 agagtgatcataacagagatcatcgattccacggtcagagtggtgagatgaaatacaggt 1080

DB 1021 agagtgatcataacagagatcatcgattccacggtcagagtggtgagatgaaatacaggt 1080

QY 1081 gaaatgtcaaacacagtttgatcacacattgtgaaagcagttctgcactgtcgttcaataat 1140
 DB 1081 gaaatgtcaaacacagtttgatcacacattgtgaaagcagttctgcactgtcgttcaataat 1140

QY 1141 ggcattgattgtgacctgtctccaagaagatcggttccattgtctgtatgagatagctcccca 1200
 DB 1141 ggcattgattgtgacctgtctccaagaagatcggttccattgtctgtatgagatagctcccca 1200

QY 1201 actgacattacccctcggagaggtgtgctgctggagacacagcgtctcattgaattgtgagac 1260
 DB 1201 actgacattacccctcggagaggtgtgctgctggagacacagcgtctcattgaattgtgagac 1260

QY 1261 ttgtatgacaagatcatgttctctgcatctctgggagatagaactataaagttatgaaacaca 1320
 DB 1261 ttgtatgacaagatcatgttctctgcatctctgggagatagaactataaagttatgaaacaca 1320

QY 1321 agtactgtgaattgtgaaagacctaataatgagacaaacagagcattgctgtttgcag 1380
 DB 1321 agtactgtgaattgtgaaagacctaataatgagacaaacagagcattgctgtttgcag 1380

QY 1381 tacagggacagcgtgtgagtgagtggtcactctgcacaacactacagattatggagacata 1440
 DB 1381 tacagggacagcgtgtgagtgagtggtcactctgcacaacactacagattatggagacata 1440

QY 1441 gaatgtgtgcatgatttaacagagctgtagaagggcactgaggaattgtgctgttattcga 1500
 DB 1441 gaatgtgtgcatgatttaacagagctgtagaagggcactgaggaattgtgctgttattcga 1500

QY 1501 ttgtatacaagagagatgagtcagtggtggtcctatgagatggaataaattgaggtgctt 1560
 DB 1501 ttgtatacaagagagatgagtcagtggtggtcctatgagatggaataaattgaggtgctt 1560

QY 1561 gtgtgtgctttggaccccgctgctcctgcagggagacactgttatacggaccccttggag 1620
 DB 1561 gtgtgtgctttggaccccgctgctcctgcagggagacactgttatacggaccccttggag 1620

QY 1621 cattccggagagatttttgcactacaagtttgatgaattccagattgcagatgttccaat 1680
 DB 1621 cattccggagagatttttgcactacaagtttgatgaattccagattgcagatgttccaat 1680

QY 1681 gatgcacaacatctctcatctggtgacttccctaataatgagtcagcggcccaagcttaaccccc 1740
 DB 1681 gatgcacaacatctctcatctggtgacttccctaataatgagtcagcggcccaagcttaaccccc 1740

QY 1741 cgttccctctctgcgaacatacacactacatctccagataaataacatacactgactcat 1800
 DB 1741 cgttccctctctgcgaacatacacactacatctccagataaataacatacactgactcat 1800

QY 1801 acttgcgccagagacccaattaaagtgtggtatattaaagttatcgtccaataccagatgagc 1860
 DB 1801 acttgcgccagagacccaattaaagtgtggtatattaaagttatcgtccaataccagatgagc 1860

QY 1861 aacaacagtaaacatacaactaacactacagttccctgcagcagcgagagcagagctt 1920
 DB 1861 aacaacagtaaacatacaactaacactacagttccctgcagcagcgagagcagagctt 1920

QY 1921 tgaagactcgttgggagacagttggtctgcagctgcggcccaagagcgttactacagacaca 1980
 DB 1921 tgaagactcgttgggagacagttggtctgcagctgcggcccaagagcgttactacagacaca 1980

QY 1981 actgactgcttaagtgctgtctatacagaagaatgcttctatacaatgtaagtatggagac 2040
 DB 1981 actgactgcttaagtgctgtctatacagaagaatgcttctatacaatgtaagtatggagac 2040

QY 2041 tttaaacctccctcctctcctccttcaactcgcagcactgattttccattgtgtc 2100
 DB 2041 tttaaacctccctcctcctcctccttcaactcgcagcactgattttccattgtgtc 2100

QY 2101 cagacaaagtgactataaataatattagtgtttgcagaaaaaaa 2151


```

Db 1139 gaaatgctaaacagcttgatccaccattgtgaagcaagcttcgacttcggttccataat 1198
QY 1141 ggcatagtatgtactgtctccaaagatcggtccattgctgtatggatataggcctccca 1200
Db 1199 ggcatagtatgtactgtctccaaagatcggtccattgctgtatggatataggcctccca 1258
QY 1201 actgaattaccctcccgaggggtgctgtgctgcagacccgagcgtcgtataagttttaac 1260
Db 1259 actgaattaccctcccgaggggtgctgtgctgcagacccgagcgtcgtataagttttaac 1318
QY 1261 ttgtatgacaagtaacattgtttctgcactcgtgggataaactataaagtaataagaca 1320
Db 1319 ttgtatgacaagtaacattgtttctgcactcgtgggataaactataaagtaataagaca 1378
QY 1321 agtactgtgaattgttaagaccttaattgacacaaagagagatctgtctgtttgacg 1380
Db 1379 agtactgtgaattgttaagaccttaattgacacaaagagagatctgtctgtttgacg 1438
QY 1381 tacagggaagagctgtgtagtgagtcacatctgacaaacacatataagataatggagaca 1440
Db 1439 tacagggaagagctgtgtagtgagtcacatctgacaaacacatataagataatggagaca 1498
QY 1441 gaattgtgtgctgtttagcagtgcttaagaagcgaatgagtgagtgctgtatctga 1500
Db 1499 gaattgtgtgctgtttagcagtgcttaagaagcgaatgagtgagtgctgtatctga 1558
QY 1501 ttgtataacaagaagatagtcagtgaggcctatgataagaaataagttgtgagatctt 1560
Db 1559 ttgtataacaagaagatagtcagtgaggcctatgataagaaataagttgtgagatctt 1618
QY 1561 gtgctgcttggagcccgctgctcctgcagagacacatctgtacagacccttgtagag 1620
Db 1619 gtgctgcttggagcccgctgctcctgcagagacacatctgtacagacccttgtagag 1678
QY 1621 catccggaaagagtttttcgcatacagtttgaattgaattcccgatgttgaagttcacat 1680
Db 1679 catccggaaagagtttttcgcatacagtttgaattgaattcccgatgttgaagttcacat 1738
QY 1681 gatgacacaatcctcatctgtgagcttccctaataatgataccagctgcgaagccccc 1740
Db 1739 gatgacacaatcctcatctgtgagcttccctaataatgataccagctgcgaagccccc 1798
QY 1741 cgttcccttccgacaatacaccatactccagataataaacaatacactgtacat 1800
Db 1799 cgttcccttccgacaatacaccatactccagataataaacaatacactgtacat 1858
QY 1801 actgcccagagaccatataag-ttgcggtatttaagctatctgcacatccacagagagag 1859
Db 1859 actgcccagagaccatataag-ttgcggtatttaagctatctgcacatccacagagagag 1918
QY 1860 caacaacagtaacaatacactactgcccagttccctgagactagccgagagagagagct 1919
Db 1919 caacaacagtaacaatacactactgcccagttccctgagactagccgagagagagagct 1978
QY 1920 ttgaagctccgttggagcacagttgtctgcaatcgcccgagagagagcttaccacagac 1979
Db 1979 ttgaagctccgttggagcacagttgtctgcaatcgcccgagagagagcttaccacagac 2038
QY 1980 aactgactgtcagtgctgtatcaagaagaatgcttcttacaatgtgaaatgttgaa 2039
Db 2039 aactgactgtcagtgctgtatcaagaagaatgcttcttacaatgtgaaatgttgaa 2098
QY 2040 cttttaaactccctccctccctcccttcaacactgcgaactagtttttccatgtgt 2099
Db 2099 cttttaaactccctccctccctcccttcaacactgcgaactagtttttccatgtgt 2158
QY 2100 ccagacaagaagtgaactataataatatttagtggtttgacagaa 2143
Db 2159 ccagacaagaagtgaactataataatatttagtggtttgacagaa 2202

```

RESULT 7
AAZ29233

```

ID AAZ29233 standard; cDNA; 2419 BP.
XX
AC AAZ29233;
XX
DT 28-FEB-2000 (first entry)
XX
DE Human cell signalling protein-12 encoding cDNA.
XX
KW Cell signalling protein-12; CSIGP-12; cell proliferation;
KW inflammatory disorder; cirrhosis; cancer; hepatitis; AIDS;
KW arteriosclerosis; Addison's disease; multiple sclerosis; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 70..1779
FT /tag= a
FT /product= "Cell Signalling Protein-12"
XX
PN WO958558-A2.
XX
PD 18-NOV-1999.
XX
PF 13-MAY-1999; 99MO-US10567.
XX
PR 13-MAY-1998; 98US-0085343.
PR 26-AUG-1998; 98US-0098010.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Hillman JL, Lal P, Yue H, Tang YT, Patterson C;
PI Baughman MR, Yang J;
XX
DR WPI: 2000-086432/07.
DR P-FSDB: AAY44249.
XX
PT Human cell signalling proteins useful for, e.g. diagnosing cell
PT proliferative and inflammatory disorders
XX
PS Claim 9; Page 87-88; 90pp; English.
XX
CC The present sequence is a cDNA obtained from Incyte clone 3239149 of
CC COLACTO1 library. It encodes cell signalling protein-12 (CSIGP-12). It
CC is expressed in musculo-skeletal, gastrointestinal and nervous
CC tissues. Fragments of CSIGP encoding nucleic acid can be used as
CC hybridisation probe for detecting CSIGP related sequences or allelic
CC variants. Recombinant CSIGP can be produced in host cells by transforming
CC them with genetically engineered vectors. Agonists or antagonists can be
CC used in the treatment of cell proliferative and inflammatory disorders
CC associated with decreased or increased CSIGP expression. CSIGP is used in
CC the diagnosis, prevention and treatment of cell proliferative disorders
CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.
XX
SQ Sequence 2419 BP; 671 A; 531 C; 625 G; 590 T; 2 other:

```

Query Match 98.3%; Score 2114.8; DB 21; Length 2419;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2119; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 11 gcgagctgacacaaaggcgcccgcgagagacgagaccagtgctcgagata 70
Db 11 ggggctgacacaaaggcgcccgcgagagacgagaccagtgctcgagata 70
QY 71 tggaccggcgagggcggtgctgcaagaagaagcactcaagtttgaattccctaaga 130
Db 71 tggaccggcgagggcggtgctgcaagaagaagcactcaagtttgaattccctaaga 130
QY 131 gagaagactgataatgtgcgaacccctaggaagataatacagaagaattcacta 190
Db 131 gagaagactgataatgtgcgaacccctaggaagataatacagaagaattcacta 190

```

OY	191	gacgacactaaacaagctctggtgccagatctcgcttaacaacgaagaacatgtttlaagca	250
Db	191	gacagacataacaacagctctgtgccagactctgcttaacaacgaagaacatgtttlaagca	250
OY	251	gacacgtcctaaagacgtgagaattgtgtggccaaaacaaactgtccaatgtccattcca	310
Db	251	gcaactgtcctaaagacgtgagaattgtgtggccaaaacaaactgtccaatgtccattcca	310
OY	311	glatatgtgtgcccaagacacacggaacatctccagccaagctctatgaaaagagaagaacgtt	370
Db	311	gtatgtatgtgtgcccaagacacacggaacatctccagccaagctctatgaaaagagaagaacgtt	370
OY	371	gtgttcaataactttggaacagtgttcagagtccagatcaataaagtgaattgtgtgacatctta	430
Db	371	gtgttcaataactttggaacagtgttcagagatccagatcaataaagtgaattgtgtgacatctta	430
OY	431	tatcccaaatgtgtcatctaccacaatgtggacataaacctcgatctctaaacctatgtttgc	490
Db	431	tatcccaaatgtgtcatctaccacaatgtggacataaacctcgatctctaaacctatgtttgc	490
OY	491	agagagatttcaataactgtctctgcagcctccagctccggagatttgatcatatgtctggaaacatcc	550
Db	491	agagagatttcaataactgtctctgcagcctccagcctccggagatttgatcatatgtctggaaacatcc	550
OY	551	tgctaatcctgtgatgtgccaaatcacatatgtgtcgtcgaactgtgtgtgcaagaaatgtcac	610
Db	551	tgctaatcctgtgatgtgccaaatcacatatgtgtcgtcgaactgtgtgtgcaagaaatgtcac	610
OY	611	gagtgaaactctgtgatgtcgtctgtgaaagaaagttaacgaagaaatgtgtcagaacaaatt	670
Db	611	gagtgaaactctgtgatgtcgtctgtgaaagaaagttaacgaagaaatgtgtcagaacaaatt	670
OY	671	ctctgtgtgaagagagcctctgcagaaacgaagaagaaatgtggagacaatattatctcaaaaacaac	730
Db	671	ctctgtgtgaagagagcctctgcagaaacgaagaagaaatgtggagacaatattatctcaaaaacaac	730
OY	731	ctctctgacgggaatgtgctctctcccaactcctttatagagcaacttaacctaaatatac	790
Db	731	ctctctgacgggaatgtgctctctcccaactcctttatagagcaacttaacctaaatatac	790
OY	791	aagacattgagaacatatgacttcaattgtggagatgtgtgaagacactgtttacacgaataatcc	850
Db	791	aagacattgagaacatatgacttcaattgtggagatgtgtgaagacactgtttacacgaataatcc	850
OY	851	actgcgaagatgtgaaacaagaacgaagatcttaacggttttaacglatgtgtatcgaagaataag	910
Db	851	actgcgaagatgtgaaacaagaacgaagatcttaacggttttaacglatgtgtatcgaagaataag	910
OY	911	taagcgcgccttcgagacacacaacatcaagaatctgtggalaataaaaacacattgtgaatgaacg	970
Db	911	taagcgcgccttcgagacacacaacatcaagaatctgtggalaataaaaacacattgtgaatgaacg	970
OY	971	gaatttccaagggcatalaagaagtttcaagttcctcgtctccaglatatgtgagaagatgtatca	1030
Db	971	gaatttccaagggcatalaagaagtttcaagttcctcgtctccaglatatgtgagaagatgtatca	1030
OY	1031	taacaggaatcatcgatctccaacgccaagtcgaagtgtgtggagttaaatcaaggttgaatgtctaa	1090
Db	1031	taacaggaatcatcgatctccaacgccaagtcgaagtgtgtggagtgttaaatcaaggttgaatgtctaa	1090
OY	1091	acacagttgatccaccatttgtgaagcagttctgtgacttggctttccataatgtgcataatgtg	1150
Db	1091	acacagttgatccaccatttgtgaagcagttctgtgacttggctttccataatgtgcataatgtg	1150
OY	1151	tgaaactgtccaaagaatcgtttccattgtcgtgtatagggaatatgtgctccccaactgacatcta	1210
Db	1151	tgaaactgtccaaagaatcgtttccattgtcgtgtatagggaatatgtgctccccaactgacatcta	1210
OY	1211	ccccccggaaggtgtcgtgtgtggaacacggaacgtctcttcaattgttlaagctttgaatga	1270
Db	1211	ccccccggaaggtgtcgtgtgtggaacacggaacgtctcttcaattgttlaagctttgaatga	1270
OY	1271	agtaacatgtttctgcatctgtgggatalagacataaagaatlaaggaaacaagaattacttgt	1330

Db	1271	agtaacatgcttcttcgcatctctggaatagaacctaaaggatctgaaacacaagtaacttg	1330
Qy	1331	aattgttaaggagaccttaaatgagacaaagaagcacttcgccgtttgcagctacaggga	1390
Db	1331	aattctgaaggaccttaaatgagacaaagaagcacttcgccgtttgcagctacaggga	1390
Qy	1391	ggctctgtagtagtgcgtccatctctacacaactacagatatactgagacatagatctg	1450
Db	1391	ggctctgtagtagtgcgtccatctctacacaactacagatatactgagacatagatctg	1450
Qy	1451	catgtctacagagtggttaagaagccacatgagaatctgctgtgatatccattgtataca	1510
Db	1451	catgtctacagagtggttaagaagccacatgagaatctgctgtgatatccattgtataca	1510
Qy	1511	agaagatagatcagtcggggccctacatgaaaaattaaatgtctgagatctctgagctgt	1570
Db	1511	agaagatagatcagtcggggccctacatgaaaaattaaatgtctgagatctctgagctgt	1570
Qy	1571	tggacccccgtgctccctcgcagaagacatctcgtctaaagacccttctgagacatccgaa	1630
Db	1571	tggacccccgtgctccctcgcagaagacatctcgtctaaagacccttctgagacatccgaa	1630
Qy	1631	gaagtttctgcagctacagtttgaatgtcaattccagattgtcagtagtltcaatgatacaca	1690
Db	1631	gaagtttctgcagctacagtttgaatgtcaattccagattgtcagtagtltcaatgatacaca	1690
Qy	1691	tccatcatctggagacttccataatgatccagctctgcacaaactgaacccccgcttccct	1750
Db	1691	tccatcatctggagacttccataatgatccagctctgcacaaactgaacccccgcttccct	1750
Qy	1751	ctcgaacacatacacctacacatctccagataaataacacatacacctgacctatacttgcag	1810
Db	1751	ctcgaacacatacacctacacatctccagataaataacacatacacctgacctatacttgcag	1810
Qy	1811	gacccattaaagttgcggtatcttaacgtatctgcacataccagatgagcaacaacagta	1870
Db	1811	gacccattaaagttgcggtatcttaacgtatctgcacataccagatgagcaacaacagta	1870
Qy	1871	aaacatcaaacacacgcagcttccctgcagctacgcagaagagcaggtttgagctct	1930
Db	1871	aaacatcaaacacacgcagcttccctgcagctacgcagaagagcaggtttgagctct	1930
Qy	1931	gtctggagacaacagttgctgcagctgcgcgcagagacagcttactcagcaaacctgact	1990
Db	1931	gtctggagacaacagttgctgcagctgcgcgcagagacagcttactcagcaaacctgact	1990
Qy	1991	tcaagtgcgtacatcagaagaatgctctctatcaatctgtgaatgtgaactttaacct	2050
Db	1991	tcaagtgcgtacatcagaagaatgctctctatcaatctgtgaatgtgaactttaacct	2050
Qy	2051	ccccctccctccctcccttaacactctgacactagtttttcccatgtgtccagaaaaag	2110
Db	2051	ccccctccctccctcccttaacactctgacactagtttttcccatgtgtccagaaaaag	2110
Qy	2111	tgaactataaatatattagtgctt	2136
Db	2111	tgaactataaatatattagtgctt	2136
RESULT 8			
AAH89966			
ID	AAH89966 standard; cDNA, 3622 BP.		
XX	AAH89966;		
AC			
XX			
XX			
DT	01-OCT-2001 (first entry)		
XX			
DE	Human bone marrow cDNA, SEQ ID NO: 97.		
KW	Human; bone marrow; antiinflammatory; cyostatic; neuroprotective;		
KW	antiviral; antibacterial; antifungal; anti-HIV; haemostatic;		
KW	immunosuppressive; gene therapy; cytokine cell proliferation;		

KM cell differentiation modulator; immune disorder; infection; cancer;
 KM human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
 OS Homo sapiens.
 XX WO200153453-A2.
 XX
 XX 26-JUL-2001.
 XX
 XX 23-DEC-2000; 2000WO-US34960.
 XX
 XX 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX
 XX WPI; 2001-488707/53.
 DR P-PSDB; AAM00847.
 XX
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 XX
 PS Claim 1; Page 274-275; 648pp; English.
 XX
 CC The present sequence is one of 251 novel human polynucleotides
 CC expressed in the bone marrow. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence
 CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 XX Sequence 3622 BP; 961 A; 839 C; 874 G; 948 T; 0 other;
 SQ

Query Match 93.5%; Score 2011.4; DB 22; Length 3622;
 Best Local Similarity 99.7%; Pred No. 0;
 Matches 2026; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 112 ttatgaattcctcagagagaagactgtlaataatgagcaacccctagaagaataata 171
 DB 261 ttccagaattcctcagagagagagactgtlaataatgagcaacccctagaagaataata 320
 QY 172 ccagagaagaatcactagacagacatacaacagctgtgcagactctgtttaacca 231
 DB 321 ccagagaagaatcactagacagacatacaacagctgtgcagactctgtttaacca 380
 QY 232 gaaacagtgatttagcaagcactgtctgaagactggaatgtgtgtgccaacaana 291
 DB 381 gaaacagtgatttagcaagcactgtctgaagactggaatgtgtgtgccaacaana 440
 QY 292 ctgtgcaatgagcactcagatgatgtgtgccaagcaaggaactcgaagaactat 351
 DB 441 ctgtgcaatgagcactcagatgatgtgtgccaagcaaggaactcgaagaactat 500
 QY 352 gaaaggaagaagactgtgttcaataactttagcagtggtcagagtcagatcaagt 411

DB 501 gaaaggaagaagactgtgttcaataactttagcagtggtcagagtcagatcaagt 560
 QY 412 gaattgtggaacacttatatacccaaatgtgtcattaccacaatgggacataactcg 471
 DB 561 gaattgtggaacacttatatacccaaatgtgtcattaccacaatgggacataactcg 620
 QY 472 tatcttaacctatgttcagagagattcataactgtctgcagctcggagattggat 511
 DB 621 tatcttaacctatgttcagagagattcataactgtctgcagctcggagattggat 680
 QY 532 catactgcaggaacactctgtcatcacttgatgccaactacatagtgtctgtaact 591
 DB 681 catactgcaggaacactctgtcatcacttgatgccaactacatagtgtctgtaact 740
 QY 592 gttgtcagaagaaatgtatccagatgactctgtagagcaatgtctgtggaagaactatcag 651
 DB 741 gttgtcagaagaaatgtatccagatgactctgtagagcaatgtctgtggaagaactatcag 800
 QY 652 agaattgtcagaagaaatgtatccagatgactctgtagagcaatgagagatgaggag 711
 DB 801 agaattgtcagaagaaatgtatccagatgactctgtagagcaatgagagatgaggag 860
 QY 712 tattattcaaaaacaaactctctgacgaggaatgtctctccaaactctttatagaga 771
 DB 861 tattattcaaaaacaaactctctgacgaggaatgtctctccaaactctttatagaga 920
 QY 772 cttatcctaaattatatacaagagacttgagagaaatagaatctaatgtgagatgtgaga 831
 DB 921 cttatcctaaattatatacaagagacttgagagaaatagaatctaatgtgagatgtgaga 980
 QY 832 catagttacagagaattcactctgcgaagtgaacaagaagaagagttactgtttacag 891
 DB 981 catagttacagagaattcactctgcgaagtgaacaagaagaagagttactgtttacag 1040
 QY 892 tatgatgatcagaagaatagttaagcgccttcgcagaaacaacaatcaagatctggagataa 951
 DB 1041 tatgatgatcagaagaatagttaagcgccttcgcagaaacaacaatcaagatctggagataa 1100
 QY 952 aaccatgtgagatcagaagcaatctcagagggccatacaagttcagtcctcttcacag 1011
 DB 1101 aaccatgtgagatcagaagcaatctcagagggccatacaagttcagtcctcttcacag 1160
 QY 1012 tatgatgagagatgatacaatacaagatcagatccagatccagatgagatgtgagatgata 1071
 DB 1161 tatgatgagagatgatacaatacaagatcagatccagatccagatgagatgtgagatgata 1220
 QY 1072 aatacaggtgaaatgatacaaacagttgattccacattgtgaaagcagttcgcagttgct 1131
 DB 1221 aatacaggtgaaatgatacaaacagttgattccacattgtgaaagcagttcgcagttgct 1280
 QY 1132 ttcaataatgcatgatgtgtacgtctccaagaatcgttccatctgtctgatatgagatg 1191
 DB 1281 ttcaataatgcatgatgtgtacgtctccaagaatcgttccatctgtctgatatgagatg 1340
 QY 1192 gctctcccaactgacatcaccctccggaagggtgtgtgtgagacaacgagctgtctcaat 1251
 DB 1341 gctctcccaactgacatcaccctccggaagggtgtgtgtgagacaacgagctgtctcaat 1400
 QY 1252 gttgtagactttagtaagagatatactgttctgcacatctggtgagatgaactataaagata 1311
 DB 1401 gttgtagactttagtaagagatatactgttctgcacatctggtgagatgaactataaagata 1460
 QY 1312 ttggaacaagaatctgtgtaattgtgtaagacctttaaattgacaacaagaagagcatgtcc 1371
 DB 1461 ttggaacaagaatctgtgtaattgtgtaagacctttaaattgacaacaagaagagcatgtcc 1520
 QY 1372 tgtttgcaagtaagagagagagctgtgtagttagtgcctatctgacaacactatcagatta 1431
 DB 1521 tgtttgcaagtaagagagagagctgtgtagttagtgcctatctgacaacactatcagatta 1580
 QY 1432 ttggaacaagaatctgtgtagttagtgcctatctgacaacactatcagatta 1491

Db	1581	tgggagacataagaatgtggtgtcgtttacagggtgtttaagaagccatbgaggaattcgtgtcg	1640
Qy	1492	tgtattcgatttgataacaaagagataatgcagctggggccattatgtgaaaaataaagctg	1551
Db	1641	tgatccgatttgataacaaagagatagtcagttggggccattatgtgaaaaataaagctg	1700
Qy	1552	tggatcttggtgtgtctttggagccccggtctccctcgaaggaacactgtctcagagacc	1611
Db	1701	tgggattcttggtgtgtctttggagccccggtctccctcgaaggaacactgtctcagagacc	1760
Qy	1612	ctgttgagacattccggaaagatttttcgactacagtttgatgaattccagattgtcagtt	1671
Db	1761	ctgttgagacattccggaaagatttttcgactacagtttgatgaattccagattgtcagtt	1820
Qy	1672	agttcaatgtatgacaaacatccctatcttgagacttccctaataatgatacagctgccagct	1731
Db	1821	agttcaatgtatgacaaacatccctatcttgagacttccctaataatgatacagctgccagct	1880
Qy	1732	gaacccccccgtctccctctctcgacatacactacatctcccgatataaataacatac	1791
Db	1881	gaacccccccgtctccctctctcgacatacactacatctcccgatataaataacatac	1940
Qy	1792	tgaacctatctgtcccaaggaaccattaaag-ctgcggtattaaagctatctgcgaat	1850
Db	1941	tgacctctactgtcccaaggaaccattaaagtttcggtattaaagctatctgcgaat	2000
Qy	1851	caggatgagcaacacagtaacaaactaacactgcgcagtttccctggagctagccag	1910
Db	2001	caggatgagcaacacagtaacaaactaacactgcgcagtttccctggagctagccag	2060
Qy	1911	agcagagctttnagaactcctgtgtggacacagttgtgtctgcagtcggccaggaagct	1870
Db	2061	agcagagctttnagaactcctgtgtggacacagttgtgtctgcagtcggccaggaagct	2120
Qy	1971	actcagcaacactgactgctctcagctgcgtcgtacatcagaagatgtctctatcaattg	2030
Db	2121	actcagcaacactgactgctctcagctgcgtcgtacatcagaagatgtctctatcaattg	2180
Qy	2031	tgattggaacttttaaacctccctccctctccctccttcaacctctgcacactagttttc	2090
Db	2181	tgattggaacttttaaacctccctccctctccctccttcaacctctgcacactagttttc	2240
Qy	2091	ccattggtctccagacaaagtgacttataatataattagtgatttcagaa	2143
Db	2241	ccattggtctccagacaaagtgacttataatataattagtgatttcagaa	2293
RESULT 9			
AAK52699			
ID	AAK52699 standard; CDNA: 3003 BP.		
XX	AAK52699;		
AC			
XX			
DT	06-NOV-2001 (first entry)		
XX			
DE	Human polynucleotide seq ID NO 2228.		
XX			
XX			
PM	WO200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US04098.		
XX			
PR	03-FEB-2000; 2000US-0496914.		
PR	27-APR-2000; 2000US-0560875.		
PR	20-JUN-2000; 2000US-0598075.		
Human; cytokine; cell proliferation; cell differentiation; gene therapy vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss.			
OS	Homo sapiens.		
XX			
XX			
XX			
XX			

Query Match	93.4%	Score 2008.2	DB 22	Length 3003
Best Local Similarity	99.6%	Pred. No. 0		
Matches 2024	Conservative	0	Mismatches 8	Indels 1
				Gaps 1
QY 112	ttatgaatcttcctcaagagagaagagctgttaataatggtgcgaacccctcggagaataata	171		
DB 298	ttccagaatctccccaagagagagagctgtataataatggtgcgaacccctcggagaataata	357		
QY 172	ccagagaagaattcactactltagacagacatacaacagctgtgcccagactctgttaaaccaa	231		
DB 358	ccagagaagaattcactactltagacagacatacaacagctgtgcccagactctgttaaaccaa	417		
QY 232	gaacacgtatgttttagcaagacactgtctatagaagacgtgaatgtgttgccaaataacaaa	291		
DB 418	gaacacgtatgttttagcaagacactgtctatagaagacgtgaatgtgttgccaaataacaaa	477		
QY 292	cttgcgaatgagcaactccacglatgactgtggtgcccacgaacggaacactccagaagctat	351		
DB 478	cttgcgaatgagcaactccacglatgactgtggtgcccacgaacggaacactccagaagctat	537		
QY 352	gaaagagaaaaggagactgtgttcaataactcttgagcagtggttccagatccagatcaagtg	411		
DB 538	gaaagagaaaaggagactgtgttcaataactcttgagcagtggttccagatccagatcaagtg	597		
QY 412	gaattgttggaacacttatcccaaatggtgcatcattacaacaatggtgcacataaactcg	471		
DB 598	gaattgttggaacacttatcccaaatggtgcatcattacaacaatggtgcacataaactcg	657		
QY 472	tatcttaaacctatgttgcagagagattccaataactgtcttcagcctcggggatttgat	531		
DB 658	tatcttaaacctatgttgcagagagattccaataactgtcttcagcctcggggatttgat	717		
QY 532	catatgtcgtgagaacatttgcataccctggtatggtcccaaacacataatgtgtgtgaact	591		
DB 718	catatgtcgtgagaacatttgcataccctggtatggtcccaaacacataatgtgtgtgaact	777		

QY 592 gtgtcagggaatggtaccgagtgacccctcgtatgcatgtgtgtggaagaagcttatccgag 651
 |||||
 Db 778 gtgtcagggaatggtaccgagtgacccctcgtatgcatgtgtgtggaagaagcttatccgag 837
 |||||
 QY 652 agaatgtcagggaatgtctctgtgtggaagagcctgtcgaagacgaagagatlggtgagag 711
 |||||
 Db 838 agaatgtcagggaatgtctctgtgtggaagagcctgtcgaagacgaagagatlggtgagag 897
 |||||
 QY 712 tattattcaaaaacaacacctctctgacgggaatgctccctcccaactctttatagaaga 771
 |||||
 Db 898 tattattcaaaaacaacacctctctgacgggaatgctccctcccaactctttatagaaga 957
 |||||
 QY 772 ctttatctcaaaaattatcaagaacatttgagacaatagaaatctaatltgagatgtggaaga 831
 |||||
 Db 958 ctttatctcaaaaattatcaagaacatttgagacaatagaaatctaatltgagatgtggaaga 1017
 |||||
 QY 832 catagtttacaagaagattctcctcgaagtgtgaaacagaagaagagtttaactgtttacaag 891
 |||||
 Db 1018 catagtttacaagaagattctcctcgaagtgtgaaacagaagaagagtttaactgtttacaag 1077
 |||||
 QY 892 tatgatgatcagaagaataagtaagcggcctctcgaagacaacacaaatcaagaatctgtggaataa 951
 |||||
 Db 1078 tatgatgatcagaagaataagtaagcggcctctcgaagacaacacaaatcaagaatctgtggaataa 1137
 |||||
 QY 952 aacacattggaatgtcagaagcgaatctcacaagccatacaggttcagttcctctgtctccag 1011
 |||||
 Db 1138 aacacattggaatgtcagaagcgaatctcacaagccatacaggttcagttcctctgtctccag 1197
 |||||
 QY 1012 tatgatgagaagaatgtatcctaagaagataatcagattccacaggttcagaagtgtggaatgta 1071
 |||||
 Db 1198 tatgatgagaagaatgtatcctaagaagataatcagattccacaggttcagaagtgtggaatgta 1257
 |||||
 QY 1072 aatcacagtgtaaatgtcctaacaacgttgatccacatttggaagcagttctgtcactgtcgt 1131
 |||||
 Db 1258 aatcacagtgtaaatgtcctaacaacgttgatccacatttggaagcagttctgtcactgtcgt 1317
 |||||
 QY 1132 ttcaataatggaatgtatgtgtgacgtgtccaaagatcgttccatgtgtatggaatgta 1191
 |||||
 Db 1318 ttcaataatggaatgtatgtgtgacgtgtccaaagatcgttccatgtgtatggaatgta 1377
 |||||
 QY 1192 gctcccccacactgacattaccctcccggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1251
 |||||
 Db 1378 gctcccccacactgacattaccctcccggaaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1437
 |||||
 QY 1252 gttgtgactttgtatgacaaatgatactgttctgtcatctgtgtgagatagaaacttaaaagta 1311
 |||||
 Db 1438 gttgtgactttgtatgacaaatgatactgttctgtcatctgtgtgagatagaaacttaaaagta 1497
 |||||
 QY 1312 tggaaacacaaatgactttgtgaaattgttaagaccttaaatgtgacacaaacgaagcatgtcc 1371
 |||||
 Db 1498 tggaaacacaaatgactttgtgaaattgttaagaccttaaatgtgacacaaacgaagcatgtcc 1557
 |||||
 QY 1372 tgtttgcagtaacagggaacggcgt 1431
 |||||
 Db 1558 tgtttgcagtaacagggaacggcgt 1617
 |||||
 QY 1432 tgggacataagaatgtgtgtgtcatgtttacgaagtgttagaagaagccatgtgaagaatgtgtgt 1491
 |||||
 Db 1618 tgggacataagaatgtgtgtgtcatgtttacgaagtgttagaagaagccatgtgaagaatgtgtgt 1677
 |||||
 QY 1492 tgtatcgaattgtataaacaagaagatagtcagtggggcctatgtalgtalgtalgtalgtalgt 1551
 |||||
 Db 1678 tgtatcgaattgtataaacaagaagatagtcagtggggcctatgtalgtalgtalgtalgtalgt 1737
 |||||
 QY 1552 tgggactctgt 1611
 |||||
 Db 1738 tgggactctgt 1797
 |||||
 QY 1612 ctgtgtgagcatctccggaagaagattttcgcatacagtttgaatgtatccaatgtcagtt 1671
 |||||
 Db 1798 ctgtgtgagcatctccggaagaagattttcgcatacagtttgaatgtatccaatgtcagtt 1857
 |||||

QY 1672 agttacatgtatgacacaatctcactgt 1731
 |||||
 Db 1858 agttacatgtatgacacaatctcactgt 1917
 |||||
 QY 1732 gaaccccccttccctctcgaacatacaccatactccagaataataaacatacac 1791
 |||||
 Db 1918 gaaccccccttccctctcgaacatacaccatactccagaataataaacatacac 1977
 |||||
 QY 1792 tgaactatcttgcacgaagccatlaaag-ttggagttatlaaglatctgtccaatc 1850
 |||||
 Db 1978 tgaactatcttgcacgaagccatlaaagtttggagttatlaaglatctgtccaatc 2037
 |||||
 QY 1851 cagagatgacacaacaagaatlaacaataactgacgacgttccctggagactgacgag 1910
 |||||
 Db 2038 cagagatgacacaacaagaatlaacaataactgacgacgttccctggagactgacgag 2097
 |||||
 QY 1911 agcagagctttggaactctgtgtgtggaacaaagtgtgtcgaagtcggccaggaagcgtct 1970
 |||||
 Db 2098 agcagagctttggaactctgtgtgtggaacaaagtgtgtcgaagtcggccaggaagcgtct 2157
 |||||
 QY 1971 actcagcacaactgactgctcagttcagttgtgtcatcagaagaatgtcttcatcaattgtga 2030
 |||||
 Db 2158 actcagcacaactgactgctcagttcagttgtgtcatcagaagaatgtcttcatcaattgtga 2217
 |||||
 QY 2031 tgaattggaactttaaacctccctcctcctccctccctccctccctccctccctccctccctccctcc 2090
 |||||
 Db 2218 tgaattggaactttaaacctccctcctcctccctccctccctccctccctccctccctccctccctcc 2277
 |||||
 QY 2091 ccattgtgtccagacaagaatgtgactttaatatattagttgttgcagaa 2143
 |||||
 Db 2278 ccattgtgtccagacaagaatgtgactttaatatattagttgttgcagaa 2330
 |||||

RESULT 10

AAK52700
 ID AAK52700 standard; cDNA; 3003 BP.

AC AAK52700;

DT 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 2229.

DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukemia;

KW nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200157190-A2.

XX 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0726422.

XX (HSE-) HYSOQ INC.

PA Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y,

PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI; 2001-476283/51.

DR P-PSDB; AAM79567.

XX Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
PS Claim 1: Page 4585-4586; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
XX Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other;

Query Match 93.4%; Score 2008.2; DB 22; Length 3003;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2024; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 112 ttatgaatccctcagagagagagacttaataatgagcccccctggagagataa 171
DB 298 ttccagaattccctcagagagagagacttaataatgagcccccctggagagataa 357
QY 172 ccagagagaatctactctagacacatacaacagctgtgcagactctgttaaccaa 231
DB 358 ccagagagaatctactctagacacatacaacagctgtgcagactctgttaaccaa 417
QY 232 gaacagatgtttaagaagcactgtatgaagactgagatgtgtgccaacaa 291
DB 418 gaacagatgtttaagaagcactgtatgaagactgagatgtgtgccaacaa 477
QY 292 ctgcgaatgcacctccagatgatgtgtcccaagcaagaaactccagcaact 351
DB 478 ctgcgaatgcacctccagatgatgtgtcccaagcaagaaactccagcaact 537
QY 352 gaaaggaagaaagactgtgtcaaatcttgagcagtggtcagagtcagatcaagt 411
DB 538 gaaaggaagaaagactgtgtcaaatcttgagcagtggtcagagtcagatcaagt 597
QY 412 gaattgtggaacatttatatcccaaatgtgcatcattacgaatgtgcaataactcg 471
DB 598 gaattgtggaacatttatatcccaaatgtgcatcattacgaatgtgcaataactcg 657
QY 472 taacttaaacctatgttcgagagagattcatalaactgtctgcagctcgagatgat 531
DB 658 taacttaaacctatgttcgagagagattcatalaactgtctgcagctcgagatgat 717
QY 532 catacgtcgaagacattctgtcaatccttgatgccaatcactatgtctgtgaact 591
DB 718 catacgtcgaagacattctgtcaatccttgatgccaatcactatgtctgtgaact 777
QY 592 gtgtgcagaagaatgtgacagagactctgagatgcatgtgtggaagaagcttaccg 651
DB 778 gtgtgcagaagaatgtgacagagactctgagatgcatgtgtggaagaagcttaccg 837
QY 652 agaatgtcagagacagattctctgtgagagagctgtgccaacgaagagatggagacg 711
DB 838 agaatgtcagagacagattctctgtgagagagctgtgccaacgaagagatggagacg 897
QY 712 taattattcaaaaacaaactctctgacggagatgtctctccaaacttttataagca 771
DB 898 taattattcaaaaacaaactctctgacggagatgtctctccaaacttttataagca 957
QY 772 cttaattcctaaattatacaagacatgagacataagaatctaatgtggagatgtggaag 831

DB 958 cttaattcctaaattatacaagacatgagacataagaatctaatgtggagatgtggaag 1017
QY 832 catagtttacagagaatctactgtccgaagtgaaacgaagcagagtttactgttcag 891
DB 1018 catagtttacagagaatctactgtccgaagtgaaacgaagcagagtttactgttcag 1077
QY 892 tatgatgctcgaagaatagtaagcggccttcgagacacaacatacaagatctggataa 951
DB 1078 tatgatgctcgaagaatagtaagcggccttcgagacacaacatacaagatctggataa 1137
QY 952 aacacatttgagatgcagagcgaattctcaacagccatagagtttactgtcttcag 1011
DB 1138 aacacatttgagatgcagagcgaattctcaacagccatagagtttactgtcttcag 1197
QY 1012 tatgtgagagatgatcatlacagagatcaggatcccaagtcgaagtgtagatga 1071
DB 1198 tatgtgagagagatgatcatlacagagatcaggatcccaagtcgaagtgtagatga 1257
QY 1072 aatacagtgaaatgtcgaacaglttgatccagatgttggaagcagttctgcactgcgt 1131
DB 1258 aatacagtgaaatgtcgaacaglttgatccagatgttggaagcagttctgcactgcgt 1317
QY 1132 ttcaataatgcatgatgtgtgacctgcgaagatcgttccattgtgtatggagatg 1191
DB 1318 ttcaataatgcatgatgtgtgacctgcgaagatcgttccattgtgtatggagatg 1377
QY 1192 gctccccaactggaatcactccctcgaaggtgtcgtgcgagacacagcgtctgcataat 1251
DB 1378 gctccccaactggaatcactccctcgaaggtgtcgtgcgagacacagcgtctgcataat 1437
QY 1252 gttgtagaactttgatgacaagatgattctctgcactgtgggataagacataaagta 1311
DB 1438 gttgtagaactttgatgacaagatgattctctgcactgtgggataagacataaagta 1497
QY 1312 tggacacaaagactgtggaattgttaagaccttaaatggaacaaacagagcattgcc 1371
DB 1498 tggacacaaagactgtggaattgttaagaccttaaatggaacaaacagagcattgcc 1557
QY 1372 tgtttcagtaagagacagcgtgtgtagtgagtggtctacatcgtgaacacatacagatta 1431
DB 1558 tgtttcagtaagagacagcgtgtgtagtgagtggtctacatcgtgaacacatacagatta 1617
QY 1432 tgggacatagaatgtgtgtcattgtaagatgttgaagagccatgtggaattgtgtcgt 1491
DB 1618 tgggacatagaatgtgtgtcattgtaagatgttgaagagccatgtggaattgtgtcgt 1677
QY 1492 tgtattcgaattgatacaacaagagatagtcagtggtggtcctatgtgtaaaatlaaagt 1551
DB 1678 tgtattcgaattgatacaacaagagatagtcagtggtggtcctatgtgtaaaatlaaagt 1737
QY 1552 tgggactctgtgctgtcttggaaccccggtcctctcaggaagacactgtctcaggaac 1611
DB 1738 tgggactctgtgctgtcttggaaccccggtcctctcaggaagacactgtctcaggaac 1797
QY 1612 ctgtgagagcattccggaagatttttgcactaagatttgatgaattccagatgtcag 1671
DB 1798 ctgtgagagcattccggaagatttttgcactaagatttgatgaattccagatgtcag 1857
QY 1672 agttcaacatgatacaacaatcctcatcctggaactcctcaatgataccagctgcgaagct 1731
DB 1858 agttcaacatgatacaacaatcctcatcctggaactcctcaatgataccagctgcgaagct 1917
QY 1732 gaaccccccggttccctctcgaacacatacactacatcctccagataaataacatacac 1791
DB 1918 gaaccccccggttccctctcgaacacatacactacatcctccagataaataacatacac 1977
QY 1792 tgaactcaactgtccgaagccattaaag-ttgcgtatttaacgatactgtccaatcac 1850
DB 1978 tgaactcaactgtccgaagccattaaag-ttgcgtatttaacgatactgtccaatcac 2037
QY 1851 caggtatgagcaacaacagtaacaatcactgtccagtttccctgagatgagcagag 1910

Db	2038	caggatgagcaacaacacgtaaacatcaactaactctgccagtttccctgactagccagag	2097
Qy	1911	agcagagctttggaactcctcgtttgggacacagtttgctcgaagtcgcacagagctct	1970
Db	2098	agcagagcctttgagactccctgcttgggacacagtttgctcgaagtcgcacagagctct	2157
Qy	1971	actagacacactgacgtcgtctgaagtcgtctgaicagagaagatgctctataacttgtaa	2050
Db	2158	actagacacactgacgtcgtctgaagtcgtctgaicacaaagatgctctataactttgttaa	2217
Qy	2031	tgattggaacttthaaacctccctcctctcctctccttcaacctgcgaactagtttttc	2090
Db	2218	tgattggaacttthaaacctccctcctcctcctctccttcaacctgcgaactagtttttc	2277
Qy	2091	ccattgtttccagacaaaagtgactataataatatttagtggttttgcacaa	2143
Db	2278	ccattgtttccagacaaaagtgactataataatatttagtggttttgcacaa	2330

	RESULT	11
AAK52701		
ID	AAK52701	standard; cDNA; 3003 BP.

AC AAK52701;

DT 06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 2230.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; antibodies; immunomodulatory; cancer; leukaemia;

Homo sapiens
05
XX

AA WO200157190-A2.
PN

PD 09-AUG-2001.

PF 05-FEB-2001;

PR 03-FEB-2000; 2000US-0496914.

PR 20-JUN-2000; 2000US-0598075

PR	01-SEP-2000; 2000US-0654936
PR	15-SEP-2000; 2000US-0654937

PR	20-OCI-2000; 2000US-06933325
PR	30-NOV-2000; 2000US-0728423

AA
PA (HYSE-) HYSEO INC.

PI Tang YT, Liu C,

Xue AJ, Yang Y, Wejhrman T, Goodrich R; PI

DR WPI; 2001-476283/51.

XX

PT useful in diagnosis and gene therapy -

PS Claim 1; Page 4586-4587; 6221pp; English.
xy

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAW78333-AAW80303) and the

cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The

polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polynucleotides have various cytokine-like activities

cc e.g. stem cell growth factor activity, haematopoiesis regulating
cc activity, tissue growth factor activity, immunomodulatory and
cc activity, tissue growth factor activity, immunomodulatory and

1

CC activating/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
XX were missing at the time of publication.
SQ Sequence 3003 BP; 807 A; 703 C; 721 G; 772 T; 0 other:

Query Match	93.48;	Score 2008.2;	DB 22;	Length 3003;
Best Local Similarity	99.68;	Pred. No. 0;		
Matches 2024; Conservative	0;	Mismatches 8;	Indels 1;	Gaps 1;

OY	112	tttltgaaatcttcacagagagagagatgtaataatgagcaaccccttggaagataa	171
Db	298	ttccagaattcttcacagagagagagactgttaataatgaggaaccccttggaagataa	357
OY	172	ccagagagaagatctactatagacagacatacaacagctgtgacagactctgtcttaacca	231
Db	358	ccagagagaagatctactatagacagacatacaacagctgtgacagactctgtcttaacca	417
OY	232	gaacagatgatttgagaagcactgcgtatgaagacctgagagattgtgtgccaacaaca	291
Db	418	gaacagatgatttgagaagcactgcgtatgaagacctgagagattgtgtgccaacaaca	477
OY	292	cttgacaaatgcaattccaaatgatgtgtcccaacaaaggaagactctcagaagcat	351
Db	478	cttgacaaatgcaattccaaatgatgtgtgtcccaacaaaggaagactctcagaagcat	537
OY	352	gaaagaagaaagagaaactgtgtgtcaaatactttgagcaagtgtgtcagaatcagataag	411
Db	538	gaaagaagaaagagaaactgtgtgtcaaatactttgagcaagtgtgtcagaatcagataag	597
OY	412	gaattgttgaaactcttatcccaagtgttcattaccacaatgaggacataaacg	471
Db	598	gaattgttgaaactcttatcccaagtgttcattaccacaatgaggacataaacg	657
OY	472	tatcttaaacctatgtgtcagagaagatcttataactgctctgcacgctcgggagattgag	531
Db	658	tatcttaaacctatgtgtcagagaagatcttataactgctctgcacgctcgggagattgag	717
OY	532	catatgcgtgagaacattctgtctatcccgtagatgccacaataactatgtgtcgtgaact	591
Db	718	catatgcgtgagaacattctgtctatcccgtagatgccacaataactatgtgtcgtgaact	777
OY	592	gtgtgcagaagaaatgtgtacccagtagactcctgtagatcatctgtgagaagaagcttaacg	651
Db	778	gtgtgcagaagaaatgtgtacccagtagactcctgtagatcatctgtgtgagaagaagcttaacg	837
OY	652	agaaatgctcagacagaattctctgtgtgagagagcctgtgcagaaacgaagagatvgggagac	711
Db	838	agaaatgctcagacagaattctctgtgtgagagagcctgtgcagaaacgaagagatvgggagac	897
OY	712	tattattcaaaaaaacaacctctctgcgggaatgcctccccaactcttttatagaca	771
Db	898	tattattcaaaaaaacaacctctctgcgggaatgcctccccaactcttttatagaca	957
OY	772	ctttatctcttaaaattatcacagacattgtgacataatgaattcaatttggagatvtgaga	831
Db	958	ctttatctcttaaaattatcacagacattgtgacataatgaattcaatttggagatvtgaga	1017
OY	832	catagtttacaagaagatctcaactgtgcgaagtgaacaaagcaagagattactgtttacag	891
Db	1018	catagtttacaagaagatctcaactgtgcgaagtgaacaaagcaagagattactgtttacag	1077
OY	892	tatgatcatcagaanaatagaaagcggcctctgagacaacaacaataagatcttggataa	951
Db	1078	tatgatcatcagaanaatagaaagcggcctctctgagacaacaacaataagatcttggataa	1137
OY	952	aacacattggaatgcagaagaaattctccacagagccatacaggtttcagtcctctgtccag	1011
Db	1138	aacacattggaatgcagaagaaattctccacagagccatacaggtttcagtcctctgtccag	1197

QY 1012 tatgatgagaaatgatacaatagatcatcgattccacggtcagagtgctggatga 1071
 Db 1198 tatgatgagaaatgatacaatagatcatcgattccacggtcagagtgctggatga 1257
 QY 1072 aatcaagtgaaatgatacaatgattcaccattgtgaagcagttctgcactgcgt 1131
 Db 1258 aatcaagtgaaatgatacaatgattcaccattgtgaagcagttctgcactgcgt 1317
 QY 1132 ttaataaagcagtgatgctgacgtccaaagatcgttcacgtgtgatggatga 1191
 Db 1318 ttaataaagcagtgatgctgacgtccaaagatcgttcacgtgtgatggatga 1377
 QY 1192 gctccccaactgacattacccctcggagggtgctgtctgcagacacggagctgtcatt 1251
 Db 1378 gctccccaactgacattacccctcggagggtgctgtctgcagacacggagctgtcatt 1437
 QY 1252 gttctgacttgatgatacaagatcattgttctgcactcggggatagaactataagata 1311
 Db 1438 gttctgacttgatgatacaagatcattgttctgcactcggggatagaactataagata 1497
 QY 1312 tggacaacaagtactgtgaattgtgaagacctaataatgacaacaaggacattgccc 1371
 Db 1498 tggacaacaagtactgtgaattgtgaagacctaataatgacaacaaggacattgccc 1557
 QY 1372 tgttgcagtlacagaggaagcgtgtagtgaatgagctcgtacatacaactacagatta 1431
 Db 1558 tgttgcagtlacagaggaagcgtgtagtgaatgagctcgtacatacaactacagatta 1617
 QY 1432 tggagacatgaatgctgtgataatgttaccagatgttgaaggccatagaaattgtgcgt 1491
 Db 1618 tggagacatgaatgctgtgataatgttaccagatgttgaaggccatagaaattgtgcgt 1677
 QY 1492 tgcattcgattgatacaagaagatgatacgtgagggccatataatgaaataatgaatga 1551
 Db 1678 tgcattcgattgatacaagaagatgatacgtgagggccatataatgaaataatgaatga 1737
 QY 1552 tggagatctgtgagctgcttggaccccgctgctcgtcaggaagacactgtctcagac 1611
 Db 1738 tggagatctgtgagctgcttggaccccgctgctcgtcaggaagacactgtctcagac 1797
 QY 1612 ctgtgagagatccggaagaatttttcgactcagtttgatgaattccagattgtcagt 1671
 Db 1798 ctgtgagagatccggaagaatttttcgactcagtttgatgaattccagattgtcagt 1857
 QY 1672 agttcaatgatgacaacaacccatcctcgtgagctcctaataatgacagctggccaaact 1731
 Db 1858 agttcaatgatgacaacaacccatcctcgtgagctcctaataatgacagctggccaaact 1917
 QY 1732 gaaccccccgtccctctcgtgaacatacaactacatcctcagataaataacatacac 1791
 Db 1918 gaaccccccgtccctctcgtgaacatacaactacatcctcagataaataacatacac 1977
 QY 1792 tgaacttatctgcccaggaacccaattaaag-cttgcggtatttaataatctcaccatac 1850
 Db 1978 tgaacttatctgcccaggaacccaattaaag-cttgcggtatttaataatctcaccatac 2037
 QY 1851 cagatgagacaacaacagatacaatacaactactgcccagttccctcggactagccgag 1910
 Db 2038 cagatgagacaacaacagatacaatacaactactgcccagttccctcggactagccgag 2097
 QY 1911 agcagagcttggagactcgttgggacacagttgctcagtcagtcgcccagagagctc 1970
 Db 2098 agcagagcttggagactcgttgggacacagttgctcagtcagtcgcccagagagctc 2157
 QY 1971 actaagacacactgagctgtcagttgctatcagaagatgcttctatcaattgtga 2030
 Db 2158 actaagacacactgagctgtcagttgctatcagaagatgcttctatcaattgtga 2217
 QY 2031 tgaattgaaactttaaaccctccctcctcctcctcctcctcctcctcctcctcctcctcctc 2090
 Db 2218 tgaattgaaactttaaaccctccctcctcctcctcctcctcctcctcctcctcctcctcctc 2277

QY 2091 coattggttcagacaagaagtgcacttataatataattagttgttccagaa 2143
 Db 2278 coattggttcagacaagaagtgcacttataatataattagttgttccagaa 2330

RESULT 12

ID AAA73132 standard; cDNA: 1707 BP.
 AC AAA73132;

DT 27-NOV-2000 (first entry)

DE Human beta-transducin repeat containing protein (beta-TrCP) cDNA.

KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; IkappaB;

KW beta-catenin; SKP1; Cull1; F-box motif; WD40 repeat motif; FMD1;

KW gene therapy; colon cancer; beta-transducin repeat containing protein;

KW beta-TrCP; ss.

OS Homo sapiens.

PN JP200016542-A.

PD 20-JUN-2000.

PF 02-DEC-1998; 98JP-0343437.

PR 02-DEC-1998; 98JP-0343437.

PA (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.

DR WPI: 2000-485550/43.

P-PSDB; AAB12813.

PT F-box protein of ubiquitin ligase SCF complex which promotes the

PT ubiquitination of IkappaB or beta-catenin

PS Disclosure: Fig 17; 19pp; Japanese.

CC The present invention describes an F-box motif protein of ubiquitin

CC ligase SCF complex which promotes the ubiquitination of IkappaB or

CC beta-catenin and is constituted by SKP1 protein, Cull1 protein and a

CC complex (SCF complex) of F-box protein containing F-box motif and Wd40

CC repeat motif and has the amino acid sequence of 43 residues (AAB12811)

CC or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin

CC ligase FMD1 protein) and (AAB12813, which is human beta-transducin

CC repeat containing protein (beta-TrCP)). The F-box protein can be used for

CC the gene therapy of colon cancer by being recombined to a virus vector.

CC The present sequence encodes the human beta-TrCP protein from the

CC present invention.

CC Sequence 1707 BP; 514 A; 349 C; 414 G; 430 T; 0 other;

SQ

Query Match 79.4%; Score 1707; DB 21; Length 1707;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 atggaccggcgagcggtgctgcaaggaagcaccacgaattatattcctcagag 129
 Db 1 atggaccggcgagcggtgctgcaaggaagcaccacgaattatattcctcagag 60
 QY 130 agagaagactgtataatgacgaacccctcctaggaagataatcagagaagaattcact 189
 Db 61 agagaagactgtataatgacgaacccctcctaggaagataatcagagaagaattcact 120
 QY 190 agacagacatacaacagctgtgcagactcgtcttaaaccaagaacagatgattga 249
 Db 121 agacagacatacaacagctgtgtccagactcgtcttaaaccaagaacagatgattga 180
 QY 250 agcactgcatataagactgagaaattgtgtggccaaacaaactggccaatggcactcc 309

Db	181	agcactgctcgtacgaagacccggaattgtgtgtgcccacaaacaaacttgcgaatgacctcc	240
QY	310	agatgatgtgtgcccagaacaggaactctcagccaagctatgaaaaagaaaagaaactg	369
Db	241	agtatgatgtgtgcccagaacaggaactctcagccaagctatgaaaaagaaaagaaactg	300
QY	370	tggtgcaataacttggagcagttgtcgagagtcagataaagttggaatttggagactctt	429
Db	301	tggtgcaataacttggagcagttgtcgagagtcagataaagttggaatttggagactctt	360
QY	430	atatcccaatgtgtcattacccaacatggtgacataaacccgcatctctaaacctatgtg	489
Db	361	atatcccaatgtgtcattacccaacatggtgacataaacccgcatctctaaacctatgtg	420
QY	490	cagaaagatttcataaactgcgtctgcagcccgaggatttgatctatatgcgttgaacatt	549
Db	421	cagaaagatttcataaactgcgtctgcagcccgaggatttgatctatatgcgttgaacatt	480
QY	550	ctgtataccttgatgtgccaatactactatgtgcgtgtbaacttgtgtgcagaagatgtac	609
Db	481	ctgtataccttgatgtgccaatactactatgtgcgtgtbaacttgtgtgcagaagatgtac	540
QY	610	cgaagtgaccttgtatgtgcatactgtgtgaaagaagcttatcagagaagaatgtgcagacagat	669
Db	541	cgaagtgaccttgtatgtgcatactgtgtgaaagaagcttatcagagaagaatgtgcagacagat	600
QY	670	tctctgtgagaagccctgtgcagcaaccgagaagagttgtggagacagatttatccaacaaa	729
Db	601	tctctgtgagaagccctgtgcagcaaccgagaagagttgtggagacagatttatccaacaaa	660
QY	730	cctcctgcagggaatgtcctctccaaactcttttatagagcacttatccataaatata	789
Db	661	cctcctgcagggaatgtcctctccaaactcttttatagagcacttatccataaatata	720
QY	790	caagacatctggagaacataagaatcacaactgtgagaatgtggaagacatagtttacaaggaatt	849
Db	721	caagacatctggagaacataagaatcacaactgtgagaatgtggaagacatagtttacaaggaatt	780
QY	850	cactgycgaagtgaaacaagcaaaaggaattactgtttacagatagatgatacagaanaata	909
Db	791	cactgycgaagtgaaacaagcaaaaggaattactgtttacagatagatgatacagaanaata	840
QY	910	gtaaagcgcccttcgagacaacacatacaaatcttggaataaaacacatctggaatgtgcag	969
Db	841	gtaaagcgcccttcgagacaacacatacaaatcttggaataaaacacatctggaatgtgcag	900
QY	970	cgaattctcacaggccatacagaattcaagtcctctgtctccagtatagatgaagagtga	1029
Db	901	cgaattctcacaggccatacagaattcaagtcctctgtctccagtatagatgaagagtga	960
QY	1030	ataacagatcatcaggattccacgcgtccagaagttgtggatcgtataacacaggtgaatactcta	1089
Db	961	ataacagatcatcaggattccacgcgtccagaagttgtggatcgtataacacaggtgaatactcta	1020
QY	1090	aacaagttgatcccaactgtgtgaagacagttctgcacattgcgtttccaataatgagcatgtg	1149
Db	1021	aacaagttgatcccaactgtgtgaagacagttctgcacattgcgtttccaataatgagcatgtg	1080
QY	1150	gtgagccgcgtcccaagaagtgttccattgtctgtatgtgagatagtgctccccaactgtacatt	1209
Db	1081	gtgagccgcgtcccaagaagtgttccattgtctgtatgtgagatagtgctccccaactgtacatt	1140
QY	1210	accctccgagaggtgtcgtgcagacacccagagctgcgtcacaagtttgtaacttgtaaac	1269
Db	1141	accctccgagaggtgtcgtgcagacacccagagctgcgtcacaagtttgtaacttgtaaac	1200
QY	1270	aagtgacattgtttctgcatactgtgtgagatagaaactataaagtatgtgacacacaagtaactgtt	1329
Db	1201	aagtgacattgtttctgcatactgtgtgagatagaaactataaagtatgtgacacacaagtaactgtt	1260
QY	1330	gaattgttgaagaccccttaaatgagacacaaaggaagcatgtgcgtgttttcagtaacagagac	1389
Db	1261	gaattgttgaagaccccttaaatgagacacaaaggaagcatgtgcgtgttttcagtaacagagac	1320

QY	1330	agcctggtgagtgagttggtcattctgtcacacactcatagattatggaacatagctgtgt	1449
Db	1321	aggctctgtagtgagttggtcattctgtcacacacactcatagattatggaacatagactgagtaagctgtgt	1380
QY	1450	gcattttacagagtggttttagaaggccatgaagaatttgctgtttatctcattgattgataac	1509
Db	1381	gcattttacagagtggttttagaaggccatgaagaatttgctgtttatctcattgattgataac	1440
QY	1510	aagagagatagtcagctggggcctctatgtatgaaaaatataaagcttgggatcttgctgtct	1569
Db	1441	aagagagatagtcagctggggcctctatgtatgaaaaatataaagcttgggatcttgctgtct	1500
QY	1570	ttggaaccccgctgtcctctcagaggaacactctgttctacggaaacctgttgagatctcgga	1629
Db	1501	ttggaaccccgctgtcctctcagaggaacactctgttctacggaaacctgttgagatctcgga	1560
QY	1630	agaatttttcgaacacaggtttgtatgtatccagattgttaagttatccatagatgacaca	1689
Db	1551	agaatttttcgaacacaggtttgtatgtatccagattgttaagttatccatagatgacaca	1620
QY	1690	atcctcatcttgagactcctccctaaatgtatccagctgtgcccagaagctcccccgcttcacct	1749
Db	1621	atcctcatcttgagactcctccctaaatgtatccagctgtgcccagaagctcccccgcttcacct	1680
QY	1750	tcttgaacaatacacactcaatctccaga	1776
Db	1681	tcttgaacaatacacactcaatctccaga	1707

RESULT	13
AAK51715	
ID	AAK51715 standard; cDNA: 2285 BP.
XX	
AC	AAK51715;
DT	06-NOV-2001 (first entry)
XX	
DE	Human polynucleotide seq ID NO 260.
XX	
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW	tissue growth factor; immunomodulatory cancer; leukaemia;
KM	nervous system disorder; arthritis; inflammation; ss.
XX	
OS	Homo sapiens.
XX	
FN	WO200157190-A2.
XX	
PD	09-AUG-2001.
XX	
PF	05-FEB-2001; 2001MO-US04098.
XX	
PR	03-FEB-2000; 2000US-0496914.
PR	27-APR-2000; 2000US-0560875.
PR	20-JUN-2000; 2000US-0598075.
PR	19-JUL-2000; 2000US-0620335.
PR	01-SEP-2000; 2000US-0654936.
PR	15-SEP-2000; 2000US-0663561.
PR	20-OCT-2000; 2000US-0693325.
PR	30-NOV-2000; 2000US-0728422.
XX	
FA	(HYSE-) HYSEQ INC.
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
PI	Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R,
XX	
DR	WPI: 2001-476283/51.
DR	P-PDB: AAM78582.
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy -

XX Claim 1: Page 1177-1180; 6221pp; English.

PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activity/inhibit activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAK80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.

XX Sequence 2285 BP; 646 A; 492 C; 573 G; 574 T; 0 other:

Query Match 78.9%; Score 1697.8; DB 22; Length 2285;

Best Local Similarity 99.0%; Pred. No. 0;
Matches 1708; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

OY 112 ttatgaattcctcagagagagagactgttaataatgagaccccttagaagataata 171
DB 277 ttcatgaattcctcagagagagagactgttaataatgagaccccttagaagataata 336
OY 172 ccgagaaagaattcattcattgagacagacatagacagctgtgacagactcgtttaaccga 231
DB 337 ccgagaaagaattcattcattgagacagacatagacagctgtgacagactcgtttaaccga 396
OY 232 gaacagctatgttttaagaagcaactgtatgaagactggaatgtgtgtgtgtgtgtgtgtgt 291
DB 397 gaacagctatgttttaagaagcaactgtatgaagactggaatgtgtgtgtgtgtgtgtgtgt 456
OY 292 ctgtgcaatgtgacactcagatgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 351
DB 457 ctgtgcaatgtgacactcagatgatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 516
OY 352 gaaagaagaaagaagaaactgt 411
DB 517 gaaagaagaaagaagaaactgt 576
OY 412 gaattgtggaacatttatatacccaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 471
DB 577 gaattgtggaacatttatatacccaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 636
OY 472 tatcttaaacctatgt 531
DB 637 tatcttaaacctatgt 696
OY 532 catatgctgtgagaaacttctgtcatatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 591
DB 697 catatgctgtgagaaacttctgtcatatccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 756
OY 592 gtgtgcaagaatgt 651
DB 757 gtgtgcaagaatgt 816
OY 652 agaatgt 711
DB 817 agaatgt 876
OY 712 tatatttcaaaaaaacttctgt 771
DB 877 tatatttcaaaaaaacttctgt 936
OY 772 cttaattcctaaattatatacaagacattgtgagacataatgtgtgtgtgtgtgtgtgtgtgt 831
DB 937 cttaattcctaaattatatacaagacattgtgagacataatgtgtgtgtgtgtgtgtgtgtgt 996

OY 832 catagttaacagaaattcactgtccgaagtgtgaaacaagcaagaggtttactgtttacag 891
DB 997 catagttaacagaaattcactgtccgaagtgtgaaacaagcaagaggtttactgtttacag 1056
OY 892 tatatgatcagaaatagtaagcgcttcctgtgagacacacaaatcaagactgtgtgtgtgtgt 951
DB 1057 tatatgatcagaaatagtaagcgcttcctgtgagacacacaaatcaagactgtgtgtgtgtgt 1116
OY 952 aacacattgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1011
DB 1117 aacacattgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1176
OY 1012 tatatgtgagaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1071
DB 1177 tatatgtgagaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1236
OY 1072 aataaagtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1131
DB 1237 aataaagtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1296
OY 1132 ttcaataatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1191
DB 1297 ttcaataatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1356
OY 1192 gcttcccaactgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1251
DB 1357 gcttcccaactgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1416
OY 1252 gt 1311
DB 1417 gt 1476
OY 1312 tggaaacaaagtaactgt 1371
DB 1477 tggaaacaaagtaactgt 1536
OY 1372 tgtttgt 1431
DB 1537 tgtttgt 1596
OY 1432 tgggacatagaaatgt 1491
DB 1597 tgggacatagaaatgt 1656
OY 1492 tgtattcgaattgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1551
DB 1657 tgtattcgaattgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1716
OY 1552 tgggattctgt 1611
DB 1717 tgggattctgt 1776
OY 1612 ctgt 1671
DB 1777 ctgt 1836
OY 1672 agttacatgtatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1731
DB 1837 agttacatgtatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaatgtgaaat 1896
OY 1732 gaacccccctgt 1791
DB 1897 gaacccccctgt 1956
OY 1792 tgaactacactgt 1836
DB 1957 tgaactacactgt 2001

RESULT 14
AAK51717
ID AAK51717 standard; CDNA; 2366 BP.
XX

QY 631 ctgtgagaagaacttatcgagagaatgtgtcagacagattctctgtgagagagccttgca 690
|||||
Db 718 ctgtgagaagaacttatcgagagaatgtgtcagacagattctctgtgagagagccttgca 777
QY 691 gaacgagagaggttggagacagatttatccaacaaacccctccgagaggaatgtcctc 750
|||||
Db 778 gaacgagagaggttggagacagatttatccaacaaacccctccgagaggaatgtcctc 837
QY 751 cccaactcttllatagagacattaloclaaaatlatacaagacattgagaacaataaa 810
|||||
Db 838 cccaactcttllatagagacattaloclaaaatlatacaagacattgagaacaataaa 897
QY 811 tctaatgtgagatgttgagagacattgtttacagagaattccatgtccgagatgaacaagc 870
|||||
Db 898 tctaatgtgagatgttgagagacattgtttacagagaattccatgtccgagatgaacaagc 957
QY 871 aaagaggttacctgttaccagatgtatgatacagaataatgaaagcgcccttcgagacaac 930
|||||
Db 958 aaagaggttacctgttaccagatgtatgatacagaataatgaaagcgcccttcgagacaac 1017
QY 931 acaataaagatctggagataaaaaacatltgaaatgcgaacgaattctcaacagccataga 990
|||||
Db 1018 acaataaagatctggagataaaaaacatltgaaatgcgaacgaattctcaacagccataga 1077
QY 991 ggttcagtcctctgtctccagatgtatgagagagtgatcataacagagatcagatcc 1050
|||||
Db 1078 ggttcagtcctctgtctccagatgtatgagagagtgatcataacagagatcagatcc 1137
QY 1051 acggttcagagtgltgagatgttaaatatacaggtgaatgtctaacaacgtltgataccaatgtc 1110
|||||
Db 1138 acggttcagagtgltgagatgttaaatatacaggtgaatgtctaacaacgtltgataccaatgtc 1197
QY 1111 gaagcgagttctgagacttgcttccaalaaatgagatgtgtgacctgtccaaagatcgt 1170
|||||
Db 1198 gaagcgagttctgagacttgcttccaalaaatgagatgtgtgacctgtccaaagatcgt 1257
QY 1171 tccatgtctgtatgagataatggcctcccaactgacattacccctccgagaggtgtcgtgc 1230
|||||
Db 1258 tccatgtctgtatgagataatggcctcccaactgacattacccctccgagaggtgtcgtgc 1317
QY 1231 ggaacccgagctgtctcaatgtgtgagacttgatgacaagtacatgttctcgtcatct 1290
|||||
Db 1318 ggaacccgagctgtctcaatgtgtgagacttgatgacaagtacatgttctcgtcatct 1377
QY 1291 ggggagatgaactataaaggtatggaacacagactgtgaaattgttaaggaccttaaat 1350
|||||
Db 1378 ggggagatgaactataaaggtatggaacacagactgtgaaattgttaaggaccttaaat 1437
QY 1351 ggaacacaaagagcatltgcctgtltgcagtacagagagcgtgtgtagtgcgtca 1410
|||||
Db 1438 ggaacacaaagagcatltgcctgtltgcagtacagagagcgtgtgtagtgcgtca 1497
QY 1411 tctgacaacactatcagattatgagacatagaatgtgtgtgcagtgttaccagatgttagaa 1470
|||||
Db 1498 tctgacaacactatcagattatgagacatagaatgtgtgtgcagtgttaccagatgttagaa 1557
QY 1471 gggcatagaggaattgtgtgcgtgtatctgcattgatacaacaagaggaatgtcagtgaggcc 1530
|||||
Db 1558 gggcatagaggaattgtgtgcgtgtatctgcattgatacaacaagaggaatgtcagtgaggcc 1617
QY 1531 tatgatvggaaaaatlaaagtvggagatctgtgcgtccttvgaaeccccgtgcctctgca 1590
|||||
Db 1618 tatgatvggaaaaatlaaagtvggagatctgtgcgtccttvgaaeccccgtgcctctgca 1677
QY 1591 gggacactctgtctacggacccctgtggagcattccggaaagatgttttgcactacagttt 1650
|||||
Db 1678 gggacactctgtctacggacccctgtggagcattccggaaagatgttttgcactacagttt 1737
QY 1651 gatgaattccagatgtgtcagatgttacaatgatgacaacaatccctcatctggagcttcccta 1710
|||||
Db 1738 gatgaattccagatgtgtcagatgttacaatgatgacaacaatccctcatctggagcttcccta 1797

QY 1711 aatgatccagctgcccgaagcttgaaaccccccggttcccttctcgaacatatacactacac 1770
|||||
Db 1798 aatgatccagctgcccgaagcttgaaaccccccggttcccttctcgaacatatacactacac 1857
QY 1771 tccagataaataacacatacactgaactataacttgcacgagacccttaagtgtgggta 1830
|||||
Db 1858 tccagataaataacacatacactgaactataacttgcacgagacccttaagtgtgggta 1917
QY 1831 tttaac 1836
|||||
Db 1918 cataac 1923

Search completed: May 11, 2002, 09:23:17
Job time: 11031 sec

Best Local Similarity 96.3%; Pred. No. 1,8e-158;
Matches 649; Conservative 0; Mismatches 23; Indels 2; Gaps 2;

```

Oy 602 aatggtaccaggtgacccctgtagcatgctgtaggaagaactatcgagagaatgta 661
    |||||||
Db 1 AATGTTACCGATGACCTCTGATGCGATGCTGTGGAAGAACCTTATCCAGAGAAATGTC 60
Oy 662 gacagatctctgtgagagagccctgacagacgaagagatgggagacagatattatca 721
    |||||||
Db 61 GGACAGATTCTCTGTGTGAGAGAGCCCTGGCAGAACGAGAGATGGGACAGATTTATTCA 120
Oy 722 aaacaaacctctgagcgggaatgctcctccaaactcttttataagacattatccta 781
    |||||||
Db 121 AAAACAAACCTCTGACCGGAGAACTCTCTCCCAACTCTTTTATAGACACTTATCTTA 180
Oy 782 aaattatcaagacattgagacatagatctaatattgagatgtggaagacatggttac 841
    |||||||
Db 181 AATATTATCAAGACATTTGAGACAAATAGATCTAATTGAGATGTGGAAGACATAGTTTAC 240
Oy 842 agagattcactgcccgaagtgaacaaagcaagagattactgttllacagatgatgac 901
    |||||||
Db 241 AGAGATTCTACTGCCGAGATGCAACCAAGAGAGTTTACTGTTTACATATCATATC 300
Oy 902 agaaatgttaagcggccttcgagacaacacatcaagatctgggataaaacacattg 961
    |||||||
Db 301 AGAAATATGTAAAGCGCCCTTCGAGACAAACACATCAAGATCTGGGATTAACACATTCG 360
Oy 962 aatgcaacgagattctcacaagacatcaggtcagttcctctgtctcagatattaga 1021
    |||||||
Db 361 AATGCAACGAAATTTCTCAGCCATACAGGTCATAGTTCCTGCTCAGATATATAGA 420
Oy 1022 gattgatacaacagagatcagatccagtcagagtgatgagatg-aaatacagat 1080
    |||||||
Db 421 GAGTATCATTAACAGATCATCGGATTCACAGGTCAGAGTGTGGATGTAATATACAGT 480
Oy 1081 gaatgcaaacacgttattcaccattgtgaacagttctgcaacttgcttcaataat 1140
    |||||||
Db 481 GAAATGCTAAACAGTTGATTTACCATTTGTAACAGTCTGCACTTGCTTCATTAATAT 540
Oy 1141 ggcattgattgtagctgtcccaagaatcgttccattgtagatgagatgagcctccca 1200
    |||||||
Db 541 GGCAATGATGTGACCTGCTCCAAAGATCGTTCCATTGCTGATGGATATGCGTCCCCCA 600
Oy 1201 actgacattacccctcggaaggtgctgctc-ggacaccgagctgctgcaatgttga 1259
    |||||||
Db 601 CTGACATTAATCCCTCGGAGGCTGCTGCGGACCAAGGCGTGTGATGTTTGA 660
Oy 1260 cttagatgacaagt 1273
    |||||
Db 661 CTTGCTTGCCGCT 674

```

RESULT 5
BE269297 692 bp mRNA linear EST 13-JUL-2000
LOCUS 601186356F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:354549 5',
DEFINITION mRNA sequence.
ACCESSION BE269297
VERSION BE269297.1 GI:9142915
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 692)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgabs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.llnl.gov
Plate: L1CM239 Row: 0 Column: 14
High quality sequence stop: 610.
Location/Qualifiers

FEATURES

1..692

source
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:354549"
/clone_11b="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph. Vector: pGB7. Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 202 a 134 c 179 g 177 t
ORIGIN

Query Match

Best Local Similarity 28.3%; Score 608.8; DB 10; Length 692;
Matches 649; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

```

Oy 602 aatggtaccaggtgacccctgtagcatgctgtaggaagaactatcgagagaatgta 661
    |||||||
Db 1 AATGTTACCGATGACCTCTGATGCGATGCTGTGGAAGAACCTTATCCAGAGAAATGTC 60
Oy 662 gacagatctctgtgagagagccctgacagacgaagagatgggagacagatattatca 721
    |||||||
Db 61 GGACAGATTCTCTGTGTGAGAGAGCCCTGGCAGAACGAGAGATGGGACAGATTTATTCA 120
Oy 722 aaacaaacctctgagcgggaatgctcctccaaactcttttataagacattatccta 781
    |||||||
Db 121 AAAACAAACCTCTGACCGGAGAACTCTCTCCCAACTCTTTTATAGACACTTATCTTA 180
Oy 782 aaattatcaagacattgagacatagatctaatattgagatgtggaagacatggttac 841
    |||||||
Db 181 AATATTATCAAGACATTTGAGACAAATAGATCTAATTGAGATGTGGAAGACATAGTTTAC 240
Oy 842 agagattcactgcccgaagtgaacaaagcaagagattactgttllacagatgatgac 901
    |||||||
Db 241 AGAGATTCTACTGCCGAGATGCAACCAAGAGAGTTTACTGTTTACATATCATATC 300
Oy 902 agaaatgttaagcggccttcgagacaacacatcaagatctgggataaaacacattg 961
    |||||||
Db 301 AGAAATATGTAAAGCGCCCTTCGAGACAAACACATCAAGATCTGGGATTAACACATTCG 360
Oy 962 aatgcaacgagattctcacaagacatcaggtcagttcctctgtctcagatattaga 1021
    |||||||
Db 361 AATGCAACGAAATTTCTCAGCCATACAGGTCATAGTTCCTGCTCAGATATATAGA 420
Oy 1022 gattgatacaacagagatcagatccagtcagagtgatgagatg-aaatacagat 1080
    |||||||
Db 421 GAGTATCATTAACAGATCATCGGATTCACAGGTCAGAGTGTGGATGTAATATACAGT 480
Oy 1081 gaatgcaaacacgttattcaccattgtgaacagttctgcaacttgcttcaataat 1141
    |||||||
Db 481 GAAATGCTAAACAGTTGATTTACCATTTGTAACAGTCTGCACTTGCTTCATTAATAT 540
Oy 1141 ggcattgattgtagctgtcccaagaatcgttccattgtagatgagatgagcctccca 1200
    |||||||
Db 541 GGCAATGATGTGACCTGCTCCAAAGATCGTTCCATTGCTGATGGATATGCGTCCCCCA 599
Oy 1201 actgacattacccctcggaaggtgctgctc-ggacaccgagctgctgcaatgttga 1260
    |||||||
Db 600 CTGACATTAATCCCTCGGAGGCTGCTGCGGAC-CCGAGCTGTGTCATAGTGTGCT 658

```

```

OY 1261 ttgtg 1264
Db 659 TTGG 662

RESULT 6
LOCUS B1822845 828 bp mRNA linear EST 04-OCT-2001
DEFINITION 603040168F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5180993 5',
mRNA sequence.
ACCESSION B1822845
VERSION B1822845.1 GI:15934395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1451 row: h column: 18
High quality sequence stop: 756.
FEATURES
source
1..828
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5180993"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH MGC Library."

BASE COUNT 241 a 187 c 215 g 185 t

Query Match 27.4%; Score 589.4; DB 10; Length 828;
Best Local Similarity 97.5%; Pred. No. 7.4e-152;
Matches 621; Conservative 0; Mismatches 11; Indels 5; Gaps 2;

OY 112 ttatgaattccccaagagaagaagactgtataatgacgaacccctaggaataata 171
Db 188 TTCCAGAAATTCCTCAGAGAGAGAAAGACTGTAAATATGCGAACCCCTAGGAAGATATA 247

OY 172 ccagagaagaattcacttagacagacatacaacagctgtgccagactctgtctaaacaa 231
Db 248 CCAGAGAAAGAAATTCATCTTAAGACAGACATACACAGCGTGCAGACTCTCTTAAACCA 307

OY 232 gaacacgatgttttagcaaacgactgtctalaagaacatgagatgtgtgtgccaacaaca 291
Db 308 GAACAGATAGTTTACGACAGACACTGCTATAGAACTCAGATTGTGTGCCCAAAACAAA 367

OY 292 ctggcaatggcacttccaglatatgtgtgtcccaagaacaggaactctcagaagaatcat 351
Db 368 CTGCCAATGGCACTTCCAGATATGTGTGCCCAAGCAACGAAACTCTCAGCAAGCTAT 427

```

```

OY 352 gaaaagaaaaggaactgtgtgtcaataactttgagcaggtgtgttagagtcagatcaagt 411
Db 428 GAAAAAGAAAAGAACTGTGTGTCAAAATCTTTGAGCAGTGTGAGTCAAGATCAAGTG 487

OY 412 gaattgtggaacatctatataccaatgtgtcattaccacaatgggcataaactcg 471
Db 488 GAATTTGTGGAACATCTTATATCCCAATGTGTCTATTAACCAACTGTGGCAGCATTAATCG 547

OY 472 tatcttaacctatgttgcagagagatttcataactgtctctgcagctcgggagttgat 531
Db 548 TATCTTAACCTATGTGTGACAGAGATTTCACTGAACCTGTGCGCAGCTCGGGGATTTGAT 607

OY 532 catatcgctggaacatctgtctatcaactgagatgccaataactatgtctgtcgtgaact 591
Db 608 CATATGCTGTGAACATCTGTCTCATACCTGATGTCGCAAAATCATATGTGCTGTGAACCT 667

OY 592 gtgtcaaggaatgtgtaccagatgacctctatgtgacatgtctgtgaaagactatcag 651
Db 668 GTGTCAAGGAATGTGTACCGAGTACCTCTGATGCGCATGCTGTGGAAGAACTTATCGAG 727

OY 652 agaatgtcagagacagatctctgtgtgagagagcctgtgcagaagaagaga---tgagg 708
Db 728 AGAATGTCAAGACAGATTTCTGTGTGAGAGGCTGCGCAGCAAGACGATGAGGAG 787

OY 709 cagtatattcaaaaacaacactctctcaggggaatg 745
Db 788 CAGTATTAATCCAAA--AAAATCTCTGACGGGAATG 822

RESULT 7
LOCUS BG829037 743 bp mRNA linear EST 22-MAY-2001
DEFINITION 602753364F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4906215 5',
mRNA sequence.
ACCESSION BG829037
VERSION BG829037.1 GI:14176624
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 743)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM1808 row: o column: 16
High quality sequence stop: 703.
FEATURES
source
1..743
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="4906215"
/clone_lib="NIH_MGC_17"
/lab_host="DH10B"
/note="Organ: muscle; Vector: pOT57; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGCAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using Zap-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT

233 a 158 c 179 g 173 t

ORIGIN

Query Match

Best Local Similarity 27.0%; Score 580.6; DB 10; Length 743;
Matches 637; Conservative 0; Mismatches 19; Indels 5; Gaps 4;

```

93 gcaagagaagagcactcaattatgaattccctcagagagagagagctgaataatggcga 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
79 GCAGAGTGGTTACTCAAAATATTATTCCTCAGAGAGAGAGAGACTGTAAATAGGCCA 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
153 accccctagaagaataatccagagaagaattcacttagacagacatacaaacgtctgc 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
139 ACCCCCTAGAGAGATATACAGAGAGAGATTCACCTTAGACAGACATACACACAGCTGTC 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
213 cagacctcttaaaccaagaacagatagtgttaagaagaagctgtatagaagctgaaga 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
199 CAGACTCTGCTTAACCAAGAAACAGTATGTTAGCAAGCACTGCTATGAAAGACTGAGAA 258
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
273 ttgtgtgcccacaaacaaactctgccaatgagcactccagatgattgtgcccagaacg 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
259 TTGTGTGGCCAAACAAACTTGCCAAATGGCAGCTTCCAGTATGATTGGCCCAAGCAACG 318
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 gaaactctagaagaagctatgaag--aaggaagaagacgtgtgtcaataactttgagcag 390
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
319 GAAACTCTCAGCAGCTATGAAACAGGAAACAGTGTGTGTCAAAATACTTTGAGCAGC 378
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
391 tggctcagagtcagatcaagtggaattgtgagacatctatacccaatgtgtcatlac 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
379 TGGTCAGAGTCAGATCAAGTGGAAATTTGTGGACATCTTATATCCCAATGTGTATTTAC 438
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
451 caacatgggacataaactcgtatcttaaacctatgttcagagaagattcataactgct 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
439 CAACTATGGGCACATTAACCTCGTATCTTAACCTATGTGCAAGAGATTTCATAACTCT 498
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
511 ctgccaagctcgggagattgatactatgcctgagaacatcttgtaacttactgtatgcca 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
499 CTGCCAGCTCGGGGATTCGATATATGCTGAGAACATTTGTCATPACCTGGATGCCAA 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
571 tcaatagtctgtcgtgaacttgctgtcagaagaaatgtlacagagtaactcctgagagcag 630
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
559 TCAGTATGTGCTGTGAACTTGCTGTGCAAGGATGTACCGAGTGAAGCTCTGATGGAGT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
631 ctgttg--aagaagctatccagagagaaatgtcagaagacatctctgtgag--aggcctgg 688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
619 CTGTGAAAGAGAGCTTATTCGAGAGATGCTCAGGACAGATTTCTGTGAGCAGGAGCCCTGG 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
688 cagaacgaaga--gagctgggagacagattattcaaaaacaacacctcctgaagagaatgct 747
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
679 CAGAAAGAAAGACGATGGGAGACAGTATTATTCCAAAAAAACCTCTGACGGGAATGCT 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 c 748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 739 c 739
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

LOCUS BG722472 752 bp mRNA linear EST 08-MAY-2001
DEFINITION 602693716F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4825970 5',
mRNA sequence.
ACCESSION BG722472
VERSION BG722472.1 GI:14001659
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 752)
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
- JOURNAL COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaabs-remail.nih.gov
Tissue Procurement: Miklos Falckovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL0739 row: p column: 03
Location/Qualifiers
1..752
High quality sequence stop: 728.

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4825970"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Organ: testis; Vector: pBluescriptR (modified
pBluescript KS+); Site:1: BamHI; Site:2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTTTT-3',
size-selected for average insert size 2.2 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT 238 a 143 c 194 g 177 t
ORIGIN

Query Match 26.3%; Score 566.4; DB 10; Length 752;
Best Local Similarity 98.8%; Pred. No. 1.7e-145;
Matches 581; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

```

115 atgaattcctcagaagagaagactgtaataatgagaaacccctagaagaataatcca 174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
161 AAGAAATTCCTCAGAGAGAGAGAACTTAATTAATGCGAACCCTTGAGAAATTAATACCA 220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 175 gagaagaattcacttagagacacatacaagaagctgtgccaagactcgttaaacgaaga 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 221 GAGAGAAATTCATTGACAGACAGATACAAAGCTGTGCCAGACTGCTTAACCAAGAA 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 235 acagatgtttagcaagacactgctatgaagactgagaatgtgtggtgccaacaaact 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 281 ACAGTATGTTTACCAAGCACTGCTATGAGAGACTGAGAAATTTGTGGCCAAAAAGAACT 340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 295 gccaatggcaactccagatgattgtgcccagaagcaagaaactcagcaagctatga 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 341 GCCAATGGCACTTCAGATATGATTGTGCCCAAGCAACGGAAACTCTCACCAAGCTATGAA 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 355 aaggaagaagaaactgtgtgcaaatccttgagcagtgctcagagtcagatcaagtgga 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 401 AAGGAAAGAGAACTGTGTGCAAAATCTTTGAGCAGTGTGACAGTCAGATCAATGAGAA 460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 415 ttgtggaacattatcccaaatgtgtcattaccacaatggtggacataaactcgtat 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 461 TTTGTGGAACATCTTATATCCCAATGTGTCAATACCAACAGGGCAACATAAATCTGTAT 520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 475 cttaaacctatgtgagagagattcataac-tgctcgaagctgggagattgagca 533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 521 CTTAAACCTATGTTGACAGAGAAATTCATTAACCTTGCTGTGCCAGCTCGGGATTGGATCA 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 534 tatcgtgagaacattctgtacatcagatgccaaatcaatcattgtctgtgaactgt 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 581 TATGTGTGAGAACATTTCTTCATTAAGCTGATGCCAAATACATATGCTGCTGAACCTGT 640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 594 gtgcagaagaatgtaccagatgacctgtgatgtgcatgtgtgtggaagaagctatagag 653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 641 GTGCAAGGAATGTACCGAGTACCTTGATGGCTGTGTGGAAGAGACTTATGAGAG 700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 654 aatgttcagagacagattctctgttgagagagcctgtgcagaagaagag 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 701 AATGTCAGACAGATTCTCTGTGGAGAGCGCTGGCAGAGACAG 748

RESULT 9
 BG518761 789 bp mRNA linear EST 02-APR-2001
 LOCUS 602578346f1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3491843 5',
 DEFINITION mRNA sequence.
 ACCESSION BG518761
 VERSION BG518761.1 GI:13513945
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 789)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM8536 row: k column: 12
 High quality sequence stop: 587.

FEATURES
 source
 1..789
 /organism="Mus musculus"
 /strain="CZECH II (fetal)"
 /db_xref="taxon:10090"
 /clone="IMAGE:3491843"
 /clone_lib="NCI_CGAP_Lu29"
 /tissue_type="spontaneous tumor, metastatic to mammary."
 /stem_cell_origin=""
 /lab_host="DH10B"
 /note="Organ: Lung; Vector: PCMV-SPORT6; Site:1; SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

BASE COUNT 257 a 176 c 202 g 154 t

ORIGIN

Query Match 24.88; Score 534.4; DB 10; Length 789;
 Best Local Similarity 86.18; Pred. No. 1,2e-136;
 Matches 661; Conservative 0; Mismatches 101; Indels 6; Gaps 6;

65 cgatattgagccggccgagcggtgctgaagaagagcaccacagttatcctc 124
 |||||
 1 CGATTATGAGACCCGCGAGCGCGTGTGCGAGAGAAAGCGCTTAATATGATTCCT 60
 |||||

125 cagagagagaagactgttaataatgagcaccocctggaagataatcacagagaagatt 184
 |||||
 61 CAGAGAGAGAAAGACTGTAAATGAATGCGAACCCTTAGAGAAATATACAGAGAAAGATT 120
 |||||

185 cacttaagagacatacaacagctgtgcagactgtccttaaacgaagaacagttatgtt 244
 |||||
 121 CACTTAAACAGACTTAAACAGCTGTGTCAGCGCTTGTGCAATTAACCAAGACAGTATGTC 180
 |||||

245 taagcaagcactgtatgaagactgagaaattgtgtgccaacaaacaaactgccaatggca 304
 |||||
 181 TAAACAAGACGCTATGAAGACTGAATAATTGTGTGGCCAAAGCAAACTTGCCCAATGGCA 240
 |||||

305 ctccagatagattgtgcccgaagcaacgaaactctcagcagcgtatgaaagaaag aag 363
 |||||
 241 CTTCACAGATGATTTGTGCCCAAGCGGAAACCTCAGCAAGAGTATGAGAAAGCAAGCAAG 300
 |||||

364 gaactgtgtcgaataactttgagcagtggtcagagtcagatcaagtgg-aatttggga 422

Db 301 GAGCTGTGTCAAGTATTGTTGAGCAGTGTGTCACAGTCTGATCAAGTGAATTTGTAGCA 360
 |||||

423 acatctatatacccaaatgtgtcattacccaacatgagcaataaactcgtatcttaacc 482
 |||||

361 ACACCTTATATCCCAAAAGTGTCTACTACACAGATGGCGACATCAACTCTTAACCTAAAC 420
 |||||

483 tatgttgagaagagatttcaataactgtctgcagccctcgaggatattgatatcgtcga 542
 |||||

421 TATGCTGAGAGGGATTTTCATTAAGTGCATCTGCAGCAGCGGGCTGTGCAACATCTGCA 480
 |||||

543 gaacattctgcataccctggaatgccaactactatgtgctgtaactgtgtgcaagga 602
 |||||

481 GAACATTCTGTCTATTAAGTGTGAGCGCAATCACTGTGCTCTCTAGCTCTGCTGCAAGGA 540
 |||||

603 atgtaccg-agtgacctgtatgagcgtgtgtggaagaagctatcagagaagatgtca 661
 |||||

541 ATGTACCGCAGATGACGTCGAGCGCATGCTGTGAAAACCTCATGCAAGAGATGCTCA 600
 |||||

662 ggaac-agattctgtgga-gagcctgtgcagacgaagagatgggagacgtattatc 719
 |||||

601 GGACGGGTACTCTGTGACAGAGCGCCGAGAGCGGAGGCTGGGACAGACTTACTC 660
 |||||

720 caaacaacacctctgcaggaatgtcctcccaactctttatagagacattatcc 779
 |||||

661 AAAACACAAACTCATGATGAGCAAGCGCTCTCCAAAGAGCAATAAGACGGCTTAAC 720
 |||||

780 taataatatacaagacattgagacaatagaaatcctaattgagagatgtg 827
 |||||

721 TAAATCATTA-TAAGACATAGAGACATTAAGTCAATGAGAAATGGGG 767
 |||||

RESULT 10
 B1103092 637 bp mRNA linear EST 26-JUN-2001
 LOCUS B1103092
 DEFINITION 602889321p1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044684
 5', mRNA sequence.
 ACCESSION B1103092
 VERSION B1103092.1 GI:14553985
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 637)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM11122 row: i column: 05
 High quality sequence stop: 629.

FEATURES
 source
 1..637
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5044684"
 /clone_lib="NCI_CGAP_Kid14"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: Kidney; Vector: PCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library. 1"

BASE COUNT 185 a 148 c 172 g 132 t


```

VERSION      B145545.1  GI:15246201
KEYWORDS
SOURCE       house mouse.
ORGANISM     Mus musculus
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE        NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL      National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT      Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cga@bbs.fmail.nih.gov
              Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LINT at:
              http://image.llnl.gov
              Plate: L14M11639 row: d column: 19
              High quality sequence start: 5
              High quality sequence stop: 684.
FEATURES
  source
    1..888
    /organism="Mus musculus"
    /strain="C57/B6"
    /db_xref="taxon:10090"
    /clone_image="5253090"
    /clone_lib="NCI_CGAP_Mam5"
    /tissue_type="tumor, gross tissue"
    /dev_stage="7 months"
    /lab_host="DH10B"
    /note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1: SalI;
    Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
    Library constructed by Life Technologies. Investigators
    providing samples: Lothar Hennighausen/Robin Humphreys,
    NIH"
BASE COUNT   219 a      218 c      248 g      203 t
ORIGIN
Query Match      24.6%; Score 528.4; DB 10; Length 888;
Best Local Similarity 81.6%; Pred. No. 6e-135;
Matches 725; Conservative 0; Mismatches 141; Indels 22; Gaps 9;
QY      854 gccgaagtgaaacaaagaggttactgttcaagatgatgcagaaatagtaa 913
DB      1 gccgaagtgaaacaaagaggttactgttcaagatgatgcagaaatagtaa 913
QY      914 gcggccttcgaagacacacatcaagatctggagataaacaacatggaatcagaacga 973
DB      61 gcggccttcgaagacacacatcaagatctggagataaacaacatggaatcagaacga 120
QY      974 ttctacagagccatcaaggttcagctctgtctccagatgatgagagatgataata 1033
DB      121 ttctacagagccatcaaggttcagctctgtctccagatgatgagagatgataata 180
QY      1034 cagagatcatcgatccacagatcagaagtgtggatgataataacagtgtaataaca 1093
DB      181 cagagatcatcgatccacagatcagaagtgtggatgataataacagtgtaataaca 240
QY      1094 cgttattcaccatctgttaa-gcagttctgcacttcggtttcaataatgcatgagtg 1152
DB      241 cgttattcaccatctgttaa-gcagttctgcacttcggtttcaataatgcatgagtg 300
QY      1153 acctgtccaaagatcgttccatctgtlatatgataatgacctcccaactgatactac 1212
DB      301 acctgtccaaagatcgttccatctgtlatatgataatgacctcccaactgatactac 360
QY      1213 ctccgagaggtgtgtgtcgaaacacgagctgtctcaatgttgtgacttgatgacaag 1272
DB      361 ctccgagaggtgtgtgtcgaaacacgagctgtctcaatgttgtgacttgatgacaag 420

```

```

QY      1273 tacatgtttctgcacatctggagtagaactataaagatgatgacaacaagactgttga 1332
DB      421 tacatgtttctgcacatctggagtagaactataaagatgatgacaacaagactgttga 480
QY      1333 ttgttaagacacctaataatgacacaagagatgctgcgtttgagtaagaagagcagg 1392
DB      481 ttgttaagacacctaataatgacacaagagatgctgcgtttgagtaagaagagcagg 540
QY      1393 ctgtgtagagtgatcattcatctgacacacatcatgataatgagacatag-aatgtgtgc 1451
DB      541 ctgtgtagagtgatcattcatctgacacacatcatgataatgagacatag-aatgtgtgc 600
QY      1452 atgttaacagagtgatgaaagccatgaaga-atgtgtgtgtgtatcattgataaac- 1509
DB      601 atgttaacagagtgatgaaagccatgaaga-atgtgtgtgtgtatcattgataaac- 660
QY      1510 aagaagatagtgatgagtgagcctatgataaataaagatgtggatc-----tgtgtg 1564
DB      661 aagaagatagtgatgagtgagcctatgataaataaagatgtggatc-----tgtgtg 720
QY      1565 ctgctttgacaccccgctgtctcctcagagagacactctgtctac-----ggaacctgtgag 1620
DB      721 ctgctttgacaccccgctgtctcctcagagagacactctgtctac-----ggaacctgtgag 780
QY      1621 c---atccggaagagttttcgactac-----agttgatgaattccag-atgtcagt 1671
DB      781 cgcattttggcgaacaaagattttccgacacccctccagttgtgtgcttcacagatgtcagc 840
QY      1672 agtccatgatgacacacatcctcatctgagacttccatgaatgatac 1719
DB      841 agtccatgatgacacacatcctcatctgagacttccatgaatgatac 888

RESULT 13
LOCUS      BG082149                      646 bp      mRNA      linear      EST 26-JAN-2001
DEFINITION H3073A06-5 N1A Mouse 15K cDNA Clone Set Mus musculus cDNA clone
ACCESSION  BG082149
VERSION     BG082149.1  GI:12564717
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE       1 (bases 1 to 646)
JOURNAL     Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka
COMMENT     Verification and initial annotation of N1A mouse 15K cDNA clone set
            Unpublished (2001)
            Other ESTs: H3073A06-3
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cda@elg.sun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3073 row: A column: 06
            Seq primer: -21M13 Reverse
            High quality sequence stop: 646
            POLYA-No.
FEATURES
  source
    1..646
    /organism="Mus musculus"
    /strain="C57BL/6J"
    /db_xref="nlaEST:H3073A06-5"
    /db_xref="taxon:10090"
    /clone="H3073A06"
    /clone_lib="N1A Mouse 15K cDNA Clone Set"
    /sex="Clones arrayed from a variety of cDNA libraries"
    /dev_stage="Clones arrayed from a variety of cDNA
    libraries"

```


Db	301	GCCAATACGAGATGAGCAACAAACACTAATCAACTACTGCCCACTTTCCTGGACT	242
Oy	1903	agccgagagacagagcgtttgagactccgtgttggaacacagttggtctcagtcgagccag	1962
Db	241	AGCCGAGAGACAGAGGCTTTGAGACTCCTCTGTGGACACAGTTGGTGTGCAATGGGCCAG	182
Oy	1963	gaacgttactcagacacactgactgcgttcagtgctgtatcagaagatgctctctaca	2022
Db	181	GAGGCTACTACGACCAACTGACTGCTTTCAGTGGCTGTATACAAAGATGCTTCAATCT	122
Oy	2023	atcgtgaatgattggaactttaaacctccctcctcctcctccttcaccltcagcaacta	2082
Db	121	TTTGTGAAATGATTGGAACCTTTAAACCTCCCTCCTCCTCCTTTCACCTCGCAACTA	62
Oy	2083	gtttttt-cccatctgtctccagacaaagtgtaactataaataataattagtgctttccag	2141
Db	61	GTTTTTTGGCCATTTGGTTCACACAAAGTACTATTAAATATTATTAAAGTTTGGCAG	2
Oy	2142	a 2142	
Db	1	A 1	

RESULT	15
LOCUS	AW762522
DEFINITION	aw762522 590 bp mRNA linear EST 04-MAY-2000 U66908.y1 NCI-CGAP_Mm3 Mus musculus cDNA clone IMAGE:3155006 5' similar to tr:Q9Z159 Q9Z159 beta-TRANSDUCIN REPEAT CONTAINING PROTEIN.; mRNA sequence.
ACCESSION	AW762522
VERSION	AW762522.1 GI:7694450
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 590)	NCI-CGAP	http://www.ncbi.nlm.nih.gov/ncicgap .		
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
	Tumor Gene Index			
	Unpublished (1997)			
	Other_ESTs: ur63608.x1			

Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov/image/html/tresources.shtml

```

MGI:1057762
Seq primer: -40RP from Gibco
High quality sequence stop: 428.
FEATURES
    Location/Qualifiers
        1..590
source

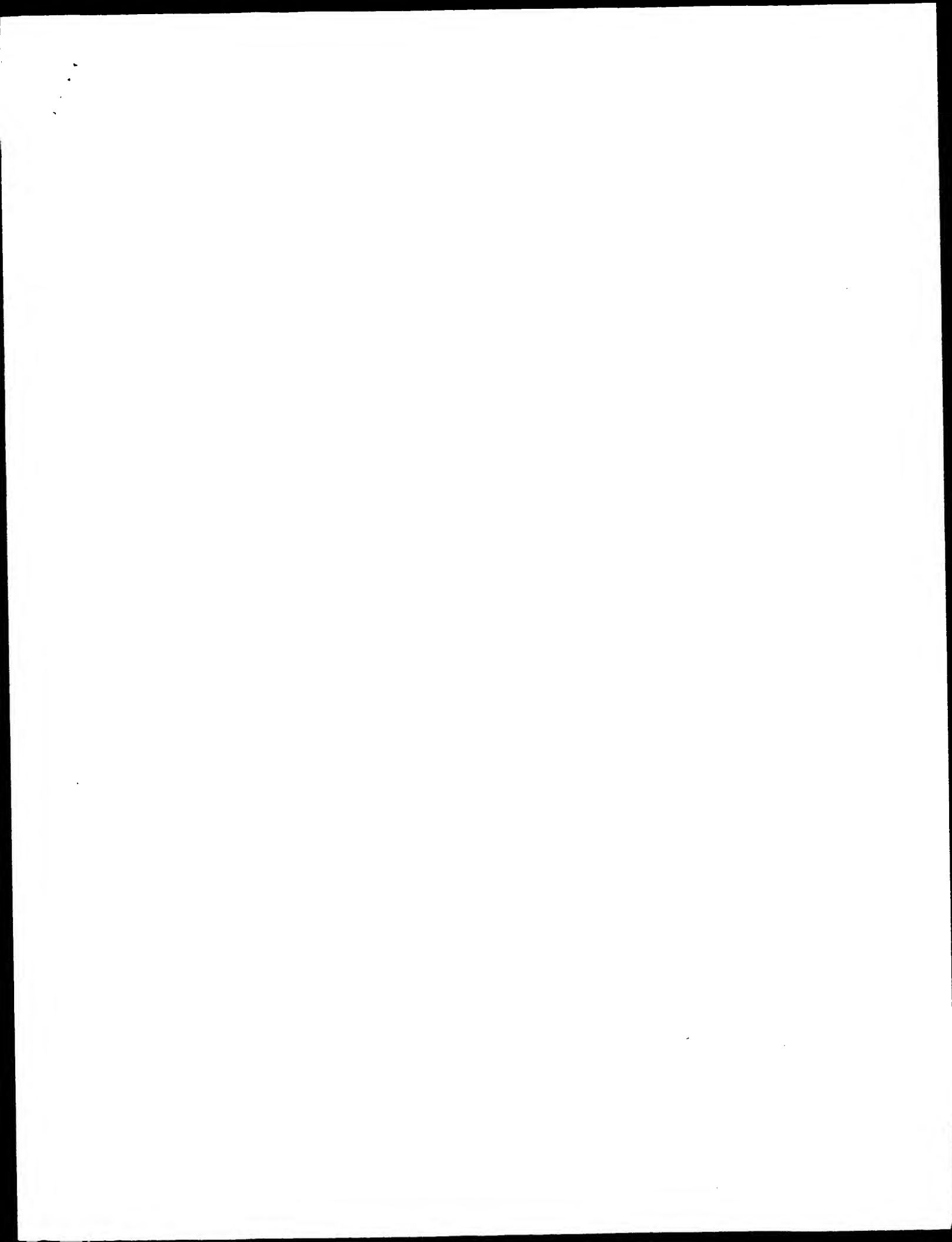
```

BASE COUNT	ORIGIN
169 a	138 c 149 g 132 t 2 others

Query Match	22.7%	Score 488.8;	DB 9;	Length 590;
Best local similarity	89.2%;	Pred. No. 4.6e-124;		
Matches 526;	Conservative	0;	Mismatches 64;	Indels 0;
				Gaps 0

QY	266	ctgaagaatttgtagggcccaaaacaaacttgccaaatgagcaatccaglatgatttgccca	32
Db	1	CTGAAATATTGTGTGGCCAAAGCCMAACTTGCCAAATGGCACTTCCAGCATGATGTGGCCA	60
QY	326	agcaacggaanaactctcagaacgctatgaaagaagaagaactgtgtgtcaaatctttg	38
Db	61	AGCAGCGGAACACTCTCAGCAGCTATGAGACAGMAAGAGAGCTGTGTCAAGTATTTTG	120
QY	386	agcaatgtgtccagatcagaatcaagtgtgaaattgtgtgaacatcttatcccaatgttc	44
Db	121	AGCAGTGGTCAAGAGTCTGTATCAAGTGAATTTGTGTGACACACTTATATCCAAATGTTC	180
QY	446	attacaacaattgggaacataaaacttgtatctctaaacctatgtctgacagaattcataa	50
Db	181	ACTACCAAGCATGGGACACATCAACTCTTCACTTAAACCATATGCTGAGAGGATTTTCATA	240
QY	506	ctgtctctgcacgctcggggatgtgatactatcgtctgagaacatctctgtcataactgtg	56
Db	241	CTGCATCGCCACACAGGGGTGTGGACACATCGCTGGAATAATTCTGTCAATCTTGGACG	300
QY	566	ccaatacactatgtgtcgtctgaacttgtgtgcaagaaatgtgtacagatgtgacctgatg	62
Db	301	CCAAATCTACTGTGTCTCTCTGTAGCTCTGTGTCCAAAGATGTGTACCGCGGTGACGTGGACG	360
QY	626	gcatalccttgagaagaagttaacgagaatattgtgtcagaacacgatctctctgtgagaagcc	68
Db	361	GCATCTGTGGAAAAAGCTCATCGAGAGAGATGGTCATMADCCGACTCTCTGTGGCGAGAGCC	420
QY	686	tggcagaacagaagagatvtggtgagcagtaatttcatcaaaaacaaacctcctgaaggaatg	74
Db	421	TGGCAGAGCCGACAGAGGCTGGGGACATATCTTATTCAAAACCAAACTCTGTATGACAGC	480
QY	746	ctctccccaactcttttatatgagcaacttatacctaataatatacaagacatltgagcaa	80
Db	481	CTCTCTCCCAACTCTTTATATGACAGCGTTTATCTTAAATCATACAGACATTGAGACAA	540
QY	806	tagaatctaaatttgagatgttgaagacataagtttacaagagaattccatgc	85
Db	541	TAGAGTCCAAATGTGAGATGTGTGGGACATATNAGTTTACAGAGATTCACATGTC	90

Search completed: May 11, 2002, 08:40:06
Job time: 13285 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:49:58 ; Search time 28.99 Seconds

(without alignments)
2870.954 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034
Sequence: 1 MDPAAVLAQKALKFNSS.....PAAQAPPSRRTYTIK 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_17.*
2: sp_archaea.*
3: sp_bacteria.*
4: sp_fungi.*
5: sp_human.*
6: sp_invertebrate.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2997	98.8	569	11 09Q0U15	09qu15 mus musculu
2	2990	98.5	569	11 092159	092159 mus musculu
3	2978	98.2	569	11 09R1G7	09r1g7 mus musculu
4	2089.5	68.9	510	5 044382	044382 drosophila
5	2089.5	68.9	510	5 09VDE3	09vde3 drosophila
6	1638.5	54.0	665	5 09GNN6	09gnn6 caenorhabd
7	1498	49.4	430	5 09B54	09b54 heterodera
8	690	22.7	506	3 09P7V1	09p7v1 schizosacch
9	661	21.8	1326	5 09V2F4	09v2f4 drosophila
10	640	21.1	553	4 09NUX6	09nux6 homo sapien
11	520	17.1	587	5 044083	044083 caenorhabd
12	387.5	12.8	942	11 09P4T2	09p4t2 mus musculu
13	365.5	12.0	942	5 096611	096611 dictyosteli
14	356	11.7	410	13 09P7R5	09p7r5 gallus gall
15	351	11.6	410	6 09GL51	09gl51 sus scrofa
16	344	11.3	411	5 096698	096698 drosophila
17	339.5	11.0	777	3 09USN3	09usn3 schizosacch
18	334.5	11.0	283	11 09JTV3	09jtv3 mus musculu
19	334.5	11.0	454	4 09HA09	09ha09 homo sapien

20	332.5	11.0	317	10 09M2Z2	09m2z2 arabidopsis
21	332	10.9	454	11 09QUH1	09quh1 mus musculu
22	330.5	10.9	594	4 09H073	09h073 homo sapien
23	330	10.9	333	10 09SY00	09sy00 arabidopsis
24	325	10.7	277	11 09R2A6	09r2a6 mus musculu
25	325	10.7	391	11 035592	035592 mus musculu
26	324.5	10.7	328	11 09D7H2	09d7h2 mus musculu
27	324.5	10.7	334	4 09NMK7	09nmk7 homo sapien
28	313.5	10.3	481	5 09VPR4	09vpr4 drosophila
29	313.5	10.3	485	4 09NVX2	09nvx2 homo sapien
30	313.5	10.3	487	4 09BU54	09bu54 homo sapien
31	312.5	10.3	330	4 09NU14	09nu14 homo sapien
32	312	10.3	876	10 09LFE2	09lfe2 arabidopsis
33	308	10.2	480	5 096995	096995 drosophila
34	304.5	10.0	1205	4 09UJ66	09uj66 homo sapien
35	304.5	10.0	1205	4 09UJ59	09uj59 homo sapien
36	301.5	9.9	476	13 093531	093531 xenopus lae
37	299.5	9.9	1205	4 09UJ65	09uj65 homo sapien
38	298.5	9.8	1171	4 09UJ58	09uj58 homo sapien
39	298.5	9.8	1205	4 09UJ67	09uj67 homo sapien
40	298.5	9.8	1205	4 09UJ64	09uj64 homo sapien
41	298.5	9.8	1205	4 09UJ63	09uj63 homo sapien
42	297	9.8	502	3 074855	074855 schizosacch
43	295.5	9.7	1205	4 09UJ60	09uj60 homo sapien
44	294.5	9.7	1249	11 09EPV5	09epv5 rattus norv
45	293.5	9.7	1205	4 09UJ62	09uj62 homo sapien

ALIGNMENTS

RESULT 1
ID 09Q0U15 PRELIMINARY; PRT; 569 AA.
AC 09Q0U15;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE UBQUITIN LIGASE FWD1 (BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN).
GN BTRC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99199275; PubMed=10097128;
RA Hakeyama S., Kitagawa M., Nakayama K., Shirane M., Matsumoto M.,
RA Hattori K., Higashi H., Nakano H., Okumura K., Onoe K., Good R.A.,
RA Nakayama K.-i.,
RT "Ubiquitin-dependent degradation of IkappaBalpha is mediated by a
RT ubiquitin ligase Skp1/Cul1/F-box protein FWD1.";
RT Proc. Natl. Acad. Sci. U.S.A. 96:3859-3863(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.,
RT "Identification of the receptor component of the IkappaBalpha-
RT ubiquitin ligase.";
RT Nature 396:590-594(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX Strausberg R.,
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF081887; AAD17755.1; -
DR EMBL: AF099932; AAD08701.1; -
DR EMBL: BC003989; AA03989.1; -
DR MGD: MGI:1338671; Btrc.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001880; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.

PRINTS: PR00320; GPROTEINBPT.
 DR SMART: SM00256; FBOX: 1.
 DR SMART: SM00320; WD40: 7.
 DR PROSITE: PS50181; FBOX: 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ligase; Repeat; WD repeat.
 SO SEQUENCE 569 AA; 65105 MW; BC7D6544815B2296 CRC64;

Query Match 98.8%; Score 2997; DB 11; Length 569;
 Best Local Similarity 98.6%; Pred. No. 5e-239;
 Matches 561; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARLCINQETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARLCINQETVCLT 60
 QY 61 STAMKTENCVAKTKLANGSSMIVPKOKRLASYEKEKELCVYFQWESDQVEVEHL 120
 DB 61 STAMKTENCVAKAKLANGSSMIVPKOKRLASYEKEKELCVYFQWESDQVEVEHL 120
 QY 121 ISQMGHOGHINSYIKPMQDPITLALPARGLDHAENILSYLDKSLCAAEIVCKEM 180
 DB 121 ISQMGHOGHINSYIKPMQDPITLALPARGLDHAENILSYLDKSLCAAEIVCKEM 180
 QY 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDGNAPNSFYRALYPKII 240
 DB 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDGNAPNSFYRALYPKII 240
 QY 241 QDIETIESNMRGRHSRLOHCRSETSGVYCYQYDOKIVSGLDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNMRGRHSRLOHCRSETSGVYCYQYDOKIVSGLDNTIKIMDKNTLECK 300
 QY 301 RLTGHTGSVYLCQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RLTGHTGSVYLCQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDKYIVSASGDRITKWNSTC 420
 DB 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDKYIVSASGDRITKWNSTC 420
 QY 421 EFWRLNGHKGRIACLOQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 480
 DB 421 EFWRLNGHKGRIACLOQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 480
 QY 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 DB 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 QY 541 ILIMDFLNDPAQAQEPSPSRITYYISR 569
 DB 541 ILIMDFLNDPAQAQEPSPSRITYYISR 569

RESULT 2
 Q92159 PRELIMINARY; PRT; 569 AA.
 AC Q92159;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE BETA-TRANSDUCIN REPEAT CONTAINING PROTEIN.
 GN BTRC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99145465; PubMed=9990853;
 RA Spencer E., Jiang J., Chen Z.J.;

RT "Signal-induced ubiquitination of IkappaBalpha by the F-box protein
 RT Slimb/Beta-Trcp.";
 RL Genes Dev. 13:284-294(1999).
 DR EMBL: AF112979; AAD04181.1; -.
 DR MGI: 1338871; Btrc-..
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBPT.
 DR SMART: SM00256; FBOX: 1.
 DR SMART: SM00320; WD40: 7.
 DR PROSITE: PS50181; FBOX: 1.
 DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SO SEQUENCE 569 AA; 65047 MW; BC7C7A44815BED96 CRC64;

Query Match 98.5%; Score 2990; DB 11; Length 569;
 Best Local Similarity 98.4%; Pred. No. 1.9e-238;
 Matches 560; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARLCINQETVCLA 60
 DB 1 MDPAEAVLOEKALKFNSSEREDCNNGEPKRIIPKNSLRQTYNSCARLCINQETVCLT 60
 QY 61 STAMKTENCVAKTKLANGSSMIVPKOKRLASYEKEKELCVYFQWESDQVEVEHL 120
 DB 61 STAMKTENCVAKAKLANGSSMIVPKOKRLASYEKEKELCVYFQWESDQVEVEHL 120
 QY 121 ISQMGHOGHINSYIKPMQDPITLALPARGLDHAENILSYLDKSLCAAEIVCKEM 180
 DB 121 ISQMGHOGHINSYIKPMQDPITLALPARGLDHAENILSYLDKSLCAAEIVCKEM 180
 QY 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDGNAPNSFYRALYPKII 240
 DB 181 RYTSOGMLMKKLIERNVRDLSMRGLAERGMGYLFKNKPPDGNAPNSFYRALYPKII 240
 QY 241 QDIETIESNMRGRHSRLOHCRSETSGVYCYQYDOKIVSGLDNTIKIMDKNTLECK 300
 DB 241 QDIETIESNMRGRHSRLOHCRSETSGVYCYQYDOKIVSGLDNTIKIMDKNTLECK 300
 QY 301 RLTGHTGSVYLCQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 RLTGHTGSVYLCQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 360
 QY 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDKYIVSASGDRITKWNSTC 420
 DB 361 VTCSKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDKYIVSASGDRITKWNSTC 420
 QY 421 EFWRLNGHKGRIACLOQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 480
 DB 421 EFWRLNGHKGRIACLOQYDERVITIGSSDSTVRWADVNGEMLNTLIHCEAVLHLRFNNGMM 480
 QY 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 DB 481 KRIVSAGVDGKIKVMDLVAALDPRAPAGTLCRTLVESHGVRFLQDFEFOIVSSSHDT 540
 QY 541 ILIMDFLNDPAQAQEPSPSRITYYISR 569
 DB 541 ILIMDFLNDPAQAQEPSPSRITYYISR 569

RESULT 3
 Q9RIG7 PRELIMINARY; PRT; 569 AA.
 AC Q9RIG7;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN.

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Winston J., Ellledge S.J., Harper J.W.;
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF110396; AAD41023.1; -
DR InterPro; IPR001810; F-box.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_6.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 569 AA; 65209 MW; E6DDCAD28D51D9D CRC64;

Query Match 98.2%; Score 2978; DB 11; Length 569;

Best Local Similarity 97.9%; Pred. No. 1.9e-237;

Matches 557; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 MDPAAEVLQEKALFKPNSREDGNGEPKRIPEKNSLRQTSYSCARLCINQETVCLIA 60
DB 1 MDPAAEVLQEKALFKPNSREDGNGEPKRIPEKNSLRQTSYSCARLCINQETVCLIT 60
QY 61 STRAKTENCVAATKLANGTSSMIVPKOKRLASYEKELCVKYEOWSESQVEFEHL 120
DB 61 STRAKTENCVAATKLANGTSSMIVPKOKRLASYEKELCVKYEOWSESQVEFEHL 120
QY 121 ISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKEY 180
DB 121 ISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKEY 180
QY 121 ISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKEY 180
DB 121 ISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKEY 180
QY 181 RYTSOGLMKKILIERVVRTDSLMRGLAERRGQYLFKNKPPDGNAPPSFYRALYPII 240
DB 181 RYTSOGLMKKILIERVVRTDSLMRGLAERRGQYLFKNKPPDGNAPPSFYRALYPII 240
QY 241 QDIETIESNWRGSRHSLORIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
DB 241 QDIETIESNWRGSRHSLORIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLECK 300
QY 301 RILTGHTGSLVLCQYDERIITIGSSDSYTVRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 360
DB 301 RILTGHTGSLVLCQYDERIITIGSSDSYTVRVMDVNTGEMLNTLIHHCCEAVLHLRNNGM 360
QY 361 VYCSKRSIAVMDMASPTDITLRVLYGHRAAVNVVDFDQKIVSASGDRITKWNSTSC 420
DB 361 VYCSKRSIAVMDMASPTDITLRVLYGHRAAVNVVDFDQKIVSASGDRITKWNSTSC 420
QY 421 ERYRTINGHKGRIACQYDRILVYSGSSDNTIRLMDIECGACRLVLEGEHELVRICIRDN 480
DB 421 ERYRTINGHKGRIACQYDRILVYSGSSDNTIRLMDIECGACRLVLEGEHELVRICIRDN 480
QY 481 KRIVSAGYGGIKIYMDLVVALDPRAPAGTLCRLTVESHGVRFLQFDFEQIVSSSHDT 540
DB 481 KRIVSAGYGGIKIYMDLVVALDPRAPAGTLCRLTVESHGVRFLQFDFEQIVSSSHDT 540
QY 541 ILIWDPLNDPAQAEPSPSRITYISR 569
DB 541 ILIWDPLNDPAQAEPSPSRITYISR 569

RESULT 4
ID 044382 PRELIMINARY: PRT: 510 AA.
AC 044382;

DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SLIMB.
GN SLIMB OR SLIMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9812115; PubMed-9461217;
RA Jüding J., Struhl G.;
RT "Regulation of the Hedgehog and Wingless signalling pathways by the F-box/WD40-repeat protein Slimb.";
RL Nature 391:493-496(1998).
DR EMBL; AF032878; AAC3852.1; -
DR FlyBase; FBgn0023423; slimb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
DR PROSITE; PS50082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; DBB0243D3730A5E8 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;

Best Local Similarity 78.3%; Pred. No. 3.4e-164;

Matches 394; Conservative 48; Mismatches 54; Indels 7; Gaps 4;

QY 64 MTEKNCVAKTLANG---TSSMIVPKOKRLAS--YEKELCVKYEOWSESQVEFE 118
DB 4 MTEKNCVAKTLANG---TSSMIVPKOKRLAS--YEKELCVKYEOWSESQVEFE 118
QY 119 HLISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKE 178
DB 63 HLISOMCHYOHGHNISYKPLMDRDFITLAPRGLDHAENILSYLAKSICAEELVCKE 178
QY 179 WYRTSDGMLMKKILIERVVRTDSLMRGLAERRGQYLFKNKPPDGNAPPSFYRALYPI 238
DB 123 WYRTSDGMLMKKILIERVVRTDSLMRGLAERRGQYLFKNKPPDGNAPPSFYRALYPI 238
QY 239 IIDIETIESNWRGSRHSLORIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 298
DB 182 IIDIETIESNWRGSRHSLORIHCRSETSKGVYCLQYDDQKIVSGLRDNTIKIWDKNTLE 298
QY 299 CKRLITGHTGSLVLCQYDERIITIGSSDSYTVRVMDVNTGEMLNTLIHHCCEAVLHLRNN 358
DB 242 CKRLITGHTGSLVLCQYDERIITIGSSDSYTVRVMDVNTGEMLNTLIHHCCEAVLHLRNN 358
QY 359 MMYTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVVDFDQKIVSASGDRITKWNSTSC 418
DB 302 MMYTCSKRSIAVMDMASPTDITLRVLYGHRAAVNVVDFDQKIVSASGDRITKWNSTSC 418
QY 419 TCERYRTINGHKGRIACQYDRILVYSGSSDNTIRLMDIECGACRLVLEGEHELVRICIR 478
DB 362 TCERYRTINGHKGRIACQYDRILVYSGSSDNTIRLMDIECGACRLVLEGEHELVRICIR 478
QY 479 DNKRIVSAGYGGIKIYMDLVVALDPRAPAGTLCRLTVESHGVRFLQFDFEQIVSSSHD 538
DB 422 DNKRIVSAGYGGIKIYMDLVVALDPRAPAGTLCRLTVESHGVRFLQFDFEQIVSSSHD 538
QY 539 DTILWDPLNDPAQAEPSPSRITYISR 561
DB 482 DTILWDPLNDPAQAEPSPSRITYISR 561

```

RESULT 5
O9VDE3 PRELIMINARY; PRT; 510 AA.
AC O9VDE3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE SLMB PROTEIN (SLMB).
GN SLMB OR CG3412.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova K.C., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Metulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny J., Moshrefi A.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reibert K., Remington K., Saunders R.D.C., Scheelers F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=OVARY AND IMAGINAL DISC;
RX MEDLINE=20245299; PubMed=10781936;
RA Miletich I., Limbourg-Bouchon B.;
RT "Drosophila null slmb clones transiently deregulate Hedgehog-
RT independent transcription of wingless in all limb discs, and induce
RT dependent transcription linked to imaginal disc regeneration.";
RL Mech. Dev. 93:15-26(2000).
DR EMBL; AE003733; AAF55853.1; -
DR EMBL; AF222924; AAF63214.1; -
DR EMBL; AF222923; AAF63213.1; -
DR FLYbase; FBgn0023423; slmb.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00646; F-box; 1.

```

```

DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR SMART; SM00256; FBOX; 1.
DR SMART; SM00320; WD40; 7.
DR PROSITE; PS50181; FBOX; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 5.
DR PROSITE; PS00082; WD_REPEATS_2; 7.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
DR Repeat; WD repeat.
DR Repeat; WD repeat.
SQ SEQUENCE 510 AA; 58952 MW; F4D5DF126F58A012 CRC64;

Query Match 68.9%; Score 2089.5; DB 5; Length 510;
Best Local Similarity 78.3%; Pred. No. 3,4e-164;
Matches 394; Conservative 49; Mismatches 53; Indels 7; Gaps 4;

QY 64 MKTECVAKTKLANG--TSSMIVPKOKLAS--YEKEKELVKYFEQWSESDQVEVE 118
DB 4 METQKIMDETIN-SNAQAFITTMIDYDPVKKDDSSPTQTERELCOFYFTQWSESGVDVE 62

QY 119 HLISQCHYQGHINSLKLMQLODFITLPAAGDHIANITISYDASLCAALYCKE 178
DB 63 HLLSRMCHYQGHQINAYLAKMLQDFITLPIGLDHIANITISYDASLKSSELVCKE 122

QY 179 WYRTSGMLMKLIEHWRTDSLMRGLARMGQYLFKNRPPDGNAPNSFYRALYK 238
DB 123 WLRITSGMLMKLIEKVTDSIMRGLARRMMQYLFKPRGQ-TQRPHSFRELPR 181

QY 239 IIDETIESNWRGRHSIORHCRSETSKGYVCLQYDQKIVSGLRDNTIKIMDKNTLE 298
DB 182 IIMNDIENNRWGRHMLRINCRESKGYVCLQYDQKIVSGLRDNTIKIMDKNTLDQ 241

QY 299 CKRLTHTGTSVLCLOYDERVITITGSSDSYRVWVDNTGEMTLTHHCEAVLHLPFNG 358
DB 242 CVKLTMGHTSVLCLOYDDKVIISGSDSTYRVWVDNTGEMVTLTHHCEAVLHLPFNG 301

QY 359 MMVTCSDRSIAVWDMSPTDITLRYLVGHRAVNVVDFDDKYIVSASGDRITIKWNTS 418
DB 302 MMVTCSDRSIAVWDMSPTSETILRYLVGHRAVNVVDFDEKXYIVSASGDRITIKWNTS 361

QY 419 TCEVTRLNHRKGIACLOYRDRLVYSSGSDNTIRIMDIECGACLVDSGHELVACIRF 478
DB 362 SCEVTRLNHRKGIACLOYRDRLVYSSGSDNIRIMDIECGACLVDSGHELVACIRF 421

QY 479 DNKRTVGAVDGKIKWMDVLAALDPRAPACTLRLTVEHSGVFRQDEPOIVSSSD 538
DB 422 DTKRTVGAVDGKIKWMDVLAALDPRASNTLCLNLTVEHSGVFRQDEPOIVSSSD 481

QY 539 DTLIMDFLNDPAAQAEPPRSPS 561
DB 482 DTLIMDFLNDPAAQAEPPRSPS 504

RESULT 6
O9GNNG PRELIMINARY; PRT; 665 AA.
AC O9GNNG;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
DE LIN-23.
GN LIN-23.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11060233;
RA Kipreos E.T., Gohel S.P., Hedgecock E.M.;
RT "The Caenorhabditis elegans F-box/WD-repeat protein LIN-23 functions
RT to limit cell division during development.";
RL Development 127:5071-5082(2000).

```


Query Match	Best Local Similarity	Matches 158; Conservative	Score 661; DB 5; Length 1326; .	34.9%; Pred. No. 9,8e-46;	76; Mismatches 167; Indels 52; Gaps 12;
105	FROMSEDOVEVEHILISOMCHYOHGHINISYKLPMLOPDTALPARAGDHTAENILSTL	164	105 FROMSEDOVEVEHILISOMCHYOHGHINISYKLPMLOPDTALPARAGDHTAENILSTL	164	105 FROMSEDOVEVEHILISOMCHYOHGHINISYKLPMLOPDTALPARAGDHTAENILSTL
852	FORMSHVEVLLADLRILDHCDPSQVRRHMKVITPEOPORFISLLRP---	907	852 FORMSHVEVLLADLRILDHCDPSQVRRHMKVITPEOPORFISLLRP---	907	852 FORMSHVEVLLADLRILDHCDPSQVRRHMKVITPEOPORFISLLRP---
165	DAKSLSAAELVCKEWRYATSDGMLMKKLEIRYVTRDTSLMKRGIAERRGNGQYLFRKKRP--	222	165 DAKSLSAAELVCKEWRYATSDGMLMKKLEIRYVTRDTSLMKRGIAERRGNGQYLFRKKRP--	222	165 DAKSLSAAELVCKEWRYATSDGMLMKKLEIRYVTRDTSLMKRGIAERRGNGQYLFRKKRP--
908	EPKDLRAAQTORSRMPLOCDDNLLMKKCRKA-----QLLAER-----SRRPRKG	953	908 EPKDLRAAQTORSRMPLOCDDNLLMKKCRKA-----QLLAER-----SRRPRKG	953	908 EPKDLRAAQTORSRMPLOCDDNLLMKKCRKA-----QLLAER-----SRRPRKG
223	-DGNAPP--NSFRALPYKTIQDIEFTIESNWCGRHSIORHCSETSKG-----VYCLQY	275	223 -DGNAPP--NSFRALPYKTIQDIEFTIESNWCGRHSIORHCSETSKG-----VYCLQY	275	223 -DGNAPP--NSFRALPYKTIQDIEFTIESNWCGRHSIORHCSETSKG-----VYCLQY
954	RDGNMPPASPAEPKAAVMR--OHI--IEMNMR-----SRPYRKPVVLGHGDHVTICLOF	1003	954 RDGNMPPASPAEPKAAVMR--OHI--IEMNMR-----SRPYRKPVVLGHGDHVTICLOF	1003	954 RDGNMPPASPAEPKAAVMR--OHI--IEMNMR-----SRPYRKPVVLGHGDHVTICLOF
276	DDOKVTSGLRBDTETIKWKNLTLECKRIILIGHGVSVLQYQDERVATITSSSDSTVAVMNVN	335	276 DDOKVTSGLRBDTETIKWKNLTLECKRIILIGHGVSVLQYQDERVATITSSSDSTVAVMNVN	335	276 DDOKVTSGLRBDTETIKWKNLTLECKRIILIGHGVSVLQYQDERVATITSSSDSTVAVMNVN
1004	SGNRIVYSSDDTTLKWSAANGKCLRTLVGHTIGVWSSQMSGNITLISDTEKTLKVMMD	1063	1004 SGNRIVYSSDDTTLKWSAANGKCLRTLVGHTIGVWSSQMSGNITLISDTEKTLKVMMD	1063	1004 SGNRIVYSSDDTTLKWSAANGKCLRTLVGHTIGVWSSQMSGNITLISDTEKTLKVMMD
336	TGEMLNTLILHHC EAVLHLRFNNGMVAWTCOSKDRSIAMVMSAPTDITLRRVYVGRAAVNV	395	336 TGEMLNTLILHHC EAVLHLRFNNGMVAWTCOSKDRSIAMVMSAPTDITLRRVYVGRAAVNV	395	336 TGEMLNTLILHHC EAVLHLRFNNGMVAWTCOSKDRSIAMVMSAPTDITLRRVYVGRAAVNV
1064	SGACVHTLQGHSTVRCHNLHGSKVSSGRDAILRGLDIEGSSCL---HVLGHIAAARC	1120	1064 SGACVHTLQGHSTVRCHNLHGSKVSSGRDAILRGLDIEGSSCL---HVLGHIAAARC	1120	1064 SGACVHTLQGHSTVRCHNLHGSKVSSGRDAILRGLDIEGSSCL---HVLGHIAAARC
396	VDPEDDKTVASAGDRTIVWNTSTCEPRTLVNGKRGKGIACIQYDRDLVYVSSDNTTFLM	455	396 VDPEDDKTVASAGDRTIVWNTSTCEPRTLVNGKRGKGIACIQYDRDLVYVSSDNTTFLM	455	396 VDPEDDKTVASAGDRTIVWNTSTCEPRTLVNGKRGKGIACIQYDRDLVYVSSDNTTFLM
1121	VQYQKTLIVSAGAYVMYKIMPEROECHTLQGHITNRYVSLQFGLIHVYSGSLPTISRW	1180	1121 VQYQKTLIVSAGAYVMYKIMPEROECHTLQGHITNRYVSLQFGLIHVYSGSLPTISRW	1180	1121 VQYQKTLIVSAGAYVMYKIMPEROECHTLQGHITNRYVSLQFGLIHVYSGSLPTISRW
456	DIEGACACIRVLEGGHEELVRCIRFDNKRIVVSGAYQCKIVMVLVAALDPRAPAGTCLNTL	515	456 DIEGACACIRVLEGGHEELVRCIRFDNKRIVVSGAYQCKIVMVLVAALDPRAPAGTCLNTL	515	456 DIEGACACIRVLEGGHEELVRCIRFDNKRIVVSGAYQCKIVMVLVAALDPRAPAGTCLNTL

Db 1181 DYETGNCCKHTLNGHOSLTSGLMELRONILVSGNADSTVAKWMDITTC-----OCLQTL 1231
 QY 516 V---EHSGRVRLQDEFOIVSSSHDDTLTMD 545
 Db 1232 SGPKNHSAVTLQFNSRFVTSDDGTIVKMD 1264

RESULT 10

Q9NUX6 PRELIMINARY; PRT; 553 AA.
 ID Q9NUX6
 AC Q9NUX6;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CDNA FL11071 FIS, CLONE PLACE1004937, MODERATELY SIMILAR TO SEL-10
 DE PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PLACENTA;
 RA Tsogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Makamatsu A.,
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.;
 RA "NDO human cDNA sequencing project."
 RL Submitted (FEBS-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK001933; BAA91986.1;
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 553 AA; 62280 MW; CA829C221986A3F2 CRC64;

Query Match 21.1%; Score 640; DB 4; Length 553;
 Best Local Similarity 30.8%; Pred. No. 1.6e-44;
 Matches 167; Conservative 92; Mismatches 211; Indels 72; Gaps 15;

QY 14 KPNSSEREDCNNGEPPRKIIPEKNSLRQTVNSCARLQNOETVCLASTAMKTENCYAKT 73
 Db 17 KLDHGEVRSFSLGKKPKCKV-----SEVYSTTGL-----VPCSA-----PTTGGDL 58
 QY 74 KLANGTSSMIYPRKGRKLSAYEKEL--CYKTFEOMSESDOVEFEHLISQMKYHGH 131
 Db 59 RAANGOG-----GQRRRTTSVQPTGLQEWLKMFSWSGGEKLLADELIDSCPTQVKH 113
 QY 132 INSYLAKPMLQDFETALPARGLDIAENILSYLDAKSICAAELVCKEMRYVTSQGMJMK 191
 Db 114 MMYVTEPQFOFDFISLP-----KELALYVSLFEPKDLQAQTCRWRIIAEDNLLMRE 169
 QY 192 LIEMVWRTSLMGLAF-----RRGWQGYLFKNKPRPDGNAPPNFYSALYPKIIIDTETI 246
 Db 170 KCKE-----EGIDBPLHIKRRK-----VIKPGFIHSPWKSAY-----IRQ--HRI 207
 QY 247 ESNMRGRHSIORIHGSEFSKGYCYQYDDQKTVSGRLRNTIKIMDKNPLECKRIITGH 306
 Db 208 DTMNRKELASPKV-LKGDHDDHTTCLOFGGNRLVSGSDNNTLKWASVAGKCLIFLVGH 266
 QY 307 TGSVLCQYDERVYITGSSSTVAVVNTGEMNLTIHHCNAVLIHLRFNGMMVYTCSD 366
 Db 267 TGVWSSQMDNIIISGSTRIKLVMAETGECIHITLYGHTSTVRCMHLHKKRVSSGS 326

QY 367 RSLAVMDASPTDITLRLVGHRAAVNVDPDKYIVSASGRTIKVMTSTCEFRIL 426
 Db 327 ATLKVMIETGQCL---HVLGHAIAVRQYDGRVRSASAYDMKVMPEIETCLHTL 383
 QY 427 NGHRGIAQLQYRDRLVYSGSSDNTIRLMDIEGACILVLEGHEELVRCIRFNKRIYSG 486
 Db 384 QGHTNRVYSLQFDGIHVHVSGLDTSIRVMDVENTGNCIHTLTGHSLSGMEKDNILVSG 443
 QY 487 AYDGKIKVMDVVALDPRAPAGTLCRTIV---EHSGRVRLQDEFOIVSSSHDDTLT 543
 Db 444 NADSTVAKIMDKTG-----OCLQTLQPGPKHOSAVTCLQFNKNEVITSSDDGTIVK 494
 QY 544 WD 545
 Db 495 WD 496

RESULT 11

Q44083 PRELIMINARY; PRT; 587 AA.
 ID Q44083
 AC Q44083;
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE SEL-10.
 GN SEL-10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-98051191; PubMed-9389650;
 RA Hubbard E.J.A., Wu G., Kitajewski J., Greenwald I.;
 RT "sel-10, a negative regulator of lin-12 activity in Caenorhabditis
 RT elegans, encodes a member of the CDC4 family of proteins."
 RL EMBL; AF020788; AAC47809.1;
 DR InterPro; IPR000412; ABC2_transport.
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PRO0320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00890; ABC2_MEMBRANE; UNKNOWN_1.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 587 AA; 65311 MW; 2D3970B4EFAA1B8C CRC64;

Query Match 17.1%; Score 520; DB 5; Length 587;
 Best Local Similarity 28.8%; Pred. No. 1.4e-34;
 Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGTSSMI-----VPRKQ--LSASYEKEL-----CYKTFEOMSESDOVE 115
 Db 35 ESSYSNGSSSSYNADKLSSRLQHKLDLSASPSRNDLNPVHEHLIALFKLSSAEDMD 94
 QY 116 FVEHLISQMKCHVHGHINSYLRMLQDFETALPARGLDIAENILSYLDAKSICAAELV 175
 Db 95 APTRLQESNMNTNIRQLAIIEPHFQDFLSCLPV-----ELGKILHNLGVDLLKVAQV 150
 QY 176 CKRWYVNTSDGMIMKL--LERV-----KTDLSMRGLAERKGGQYLFKNKPRPDGNAP 227
 Db 151 SKWKVLITSEDKIMKSLGVEEFKHHDPDTRVGTAGMOGTALNAG-----VTIPDHIOP 203
 QY 228 PN-SFYRALYPRIIODI-----ETIESNMRGRHSIORIHCRSPTSFGYCL 273

Db 204 CDLWHPFLKLOKEDIFERAADSKRYLRADKIKENNNANPIMSGAV-LRGHEHDVITCM 262
 Qy 274 QYDDKIVSGLRDNTIKIMDKNTLECKRILGHTGSLCLOYDE--RVIIITGSSDSSTVRY 331
 Db 263 QIHDDVLVTGSDDDNTLKWKYIDKGEVMTYLVGHTGVTWSQISOCGRYIVSGSDTRIVKV 322
 Qy 332 MDVNTGEMLNLIHHCCEAVLHFRNGNMVYCSKDRSLAVYMDMASPDIITLRVLYGHR 391
 Db 333 WSTVDSLSLHTLQGHSTVRCMAMAGSLTVGSRDTLLRVMDVESGRHLA---TLHGHH 379
 Qy 392 AVNVDFDXYIVSASGDRITKVMNTSTCEPFTLNGHKGRIACLOYRDR--LVVSGSSD 449
 Db 380 AVRCQVPGTIVTSGGYPTFKINNAHGRICRITLTGNHNVYSLLESEHSYVCSGLD 439
 Qy 450 NTRILMDI--ECGACLVRLGHELVRCIRPDNKRIVSGYDGKIKVMDIYVALDPRAP 506
 Db 440 TSIKRWDETRPGECEVALLQGHSTSLTSGMOLRGNIILVSCNASHVHWDT-----H 491
 Qy 507 AGTLCRLTVHSGRVFRLQ--FDEFOIVSSSHDITLLMD 545
 Db 492 EGT-CVHMLSGHRSATISLQWFGNNMVAITSSDDGVTKLMD 530

RESULT 12

Q9D4T2 PRELIMINARY: PRT: 304 AA.
 ID 09D4T2
 AC 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 4930563E19RIK PROTEIN.
 GN 4930563E19RIK
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Toh M., Ishii Y., Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I., Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H., Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombert P., Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F., Suzuki H., Toyokawa K., Wang K., Weitz C., Whitaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohlsuki S., Hayashizaki Y.;
 RA "functional annotation of a full-length mouse cDNA collection."
 RT Nature 409:685-690(2001).
 RL EMBL; AK016201; BAB30146.1;
 DR MGD; MGI:1923089; 4930563E19RIK.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE; PS0082; WD_REPEATS_2; 7.
 DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
 KW Repeat: WD repeat.
 SQ SEQUENCE 304 AA; 33173 MW; AFAACDC487939E3D CRC64;

Query Match 12.8%; Score 387.5; Db 11; Length 304;
 Best Local Similarity 27.8%; Pred. No. 5e-24;
 Matches 84; Conservative 72; Mismatches 119; Indels 27; Gaps 9;
 Qy 258 QRIHCSESTSKGYVCLQYDD--OKIVSGLRDNTIKIMDKNTLECKRILGHTGSLCLOY 314
 Db 14 EELHTLEGHKNVYVYALFANNPYGDKIATGSPDKTCLMGAELGKCYHFRGHTAIVCL 73
 Qy 315 YDNR--VIITGSSDSVVRVWNTVGMNLTLHHCCEAVLHFRNGN--MMVYCSKDRSLA 370
 Db 74 FNPSTVAVAGSDTAKLMDIONGEEVYTLGHAEIISLFSFDSIGRITIGSDPHYV 133
 Qy 371 VMDASPTDITLRVLYGHRAAVN--VDEDDKYIVSASGDRITKVMNTSTCEPFTLNG 428
 Db 134 VMDASGKRV---HTLGHCAEISSALFMMDCSLITGSDMTCLMDATSGKYAVALTLG 190
 Qy 429 HRRGT--ACLOYDRILVSGSSDNTIRLMDIEGACLRVLEHHELVRCIRRD--NKRIV 484
 Db 191 HDEIILDSCEPDYVGLKILAVASADGTARYVNTTRKCVTLKEGHEISKISFNPGNRIL 250
 Qy 485 SGAYDGIKIVMDIYVALDPRAPAGTLCRLTVHSGRVFRLQFDEFO--IVSSSHDITLL 542
 Db 251 TGSSDKTARIMDVQTC-----QCLQVLESHTEIFSCAFNYKGNIVITGSDNDSR 301
 Qy 543 IW 544
 Db 302 IW 303

RESULT 13

Q96611 PRELIMINARY: PRT: 942 AA.
 ID 096611
 AC 096611
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MEK KINASE ALPHA.
 GN MKKA.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
 NX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KAX-3;
 RX MEDLINE=99051319; PubMed=9832508;
 RA Chung C.Y., Reddy T.B.K., Zhou K., Fitrel R.A.;
 RT "A novel, putative MEK kinase controls developmental timing and spatial patterning in Dictyostelium and is regulated by ubiquitin-mediated protein degradation";
 RT Genes Dev. 12:3564-3578(1998).
 RL -i- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF093689; AAC97114.1;
 DR InterPro: IPR000719; Euk-kinase.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR000270; OPR.
 DR InterPro: IPR002290; Ser_thr_kin_actsite.
 DR InterPro: IPR001245; Tyr_kin.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR PRINTS; PR00109; TYRKINASE.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00016; OPR; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS0181; FBOX; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_4.
 DR PROSITE; PS0082; WD_REPEATS_2; 5.

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KM ATP-binding; Kinase; Repeat; Serine/threonine-protein kinase;
 KW Transferrase; WD repeat.
 SQ SEQUENCE 942 AA; 105796 MW; C9E4928A8C7C68F7 CRC64;

Query Match 12.0%; Score 365.5; DB 5; Length 942;
 Best Local Similarity 24.7%; Pred. No. 1.5e-21;
 Matches 110; Conservative 67; Mismatches 140; Indels 129; Gaps 16;

QY 175 VCKEYRVTSYDGMIMKK-----LIERVVRTDSL-WRG-----LAERRGMOQYLFKKPPD 223
 DB 547 VCKHMOIIDDDELMMKCYSDRLINKSFESTIMKSNYIKIKQKW----FHKKL-- 600
 QY 224 GMAPNSFYALPKIIIDLETIESMRCGRSLQRIHCRSETSKGVYCLQ-YDDQ-KIV 281
 DB 601 -----NHSTLNGH-----DKGVFCVKLIDDOGMVL 625
 QY 282 SGLRNTIKIMD-----KNTLECKR----- 301
 DB 626 SGGEDKKLKVMYDISGHHNHSGIVGSIKSKGLIINNNSNNSSSSSSSSRL 685
 QY 302 -ILTGTSVLCIQ-----DERVITGSSDSYRVWVDVNTGEMLNTLIHCEAVLHR 354
 DB 686 FSLKSHSGCIKSYDQYORSGSDVSRVFETASADFTCKIFSLTKKTLFTYTNHQAATCIN 745
 QY 355 FNNGM---WVTCSDRSIAVWDMASPTDITLRVVGHRAAVNVDF-----DKY 402
 DB 746 YLGDVANKCITSSLDKTIQMDAETGSCSLTLR--GHGGIYCVKTDVATNGNGVNL 802
 QY 403 IYASGDRITKIVWNTSTCEVFTLNGHKRGIAQLQYRDLVYVSGSSDNTIRLMDIECGAC 462
 DB 803 IYASVADKTSNWDTRSSSKVRSFOHTEDVLCVYFDQKVYTGSCDGIKIMDITGKT 862
 QY 463 LNVLEGHE---ELVRCIFEDKRIYSGAYDKIKVMDLVAALDPRAPAGTICLTIVER 518
 DB 863 ISTFLPSETRKQNYVYVDFDQSKIISGKTIIRIMDIYNERDSRISIG-----H 913
 QY 519 SGRVRLQDFEFOIYSSSHDDTILM 544
 DB 914 HETIFSLQFNNOKLITGSLDKLVKIM 939
 RESULT 14
 Q9PTR5 PRELIMINARY; PRT; 410 AA.
 AC Q9PTR5; 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 GN L1S1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus
 NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Reiner O., Shmueli O.;
 RT "Characterization of the chicken homolog of L1S1";
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113946; AAF18938.1; -
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_5.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 410 AA; 46664 MW; FCS584BD06EDCA20 CRC64;

Query Match 11.7%; Score 356; DB 13; Length 410;
 Best Local Similarity 29.5%; Pred. No. 3e-21;
 Matches 88; Conservative 56; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLRDNTIKIMDNTLECKRILTGHTGSLVCLQYDE--RIYITGSSDSYRVWVDVNTG 337
 DB 123 IYASASDATIKVMYDETQDFERTLKGHTDSYODISFDHTGKLKILASCSADMRTIKLMDQGF 182
 QY 338 EMLNTLIHCEAVLHRF--NNGMAYTCSKDRSIAVWDMASPTDITLRVVGHRAAVNV 395
 DB 183 ECIIRTMGHGDNHVSVAIMPNDHIVASARDKTIKWEVQTYCV---KTFGRHREVRM 239
 QY 396 V--DFDDKIYVASGDRITKIVWNTSTCEVFTLNGHKRGIAQLQYRDR----- 441
 DB 240 VRPNQDGLILASCSNDQYRVWVAVATKECKALREHENVYECISWAPESYSTSEATGS 299
 QY 442 -----LVYSSSDNTIRLMDIECGACLRVLEGHEELVRCIRPDN--KRIYSGAYD 489
 DB 300 ETKKSGKRPPELLSGSRDRTIKIMDISGMCLMTLVGHDMVRGVLFHSGCKFILSCAD 359
 QY 490 GKIKYMDLVAALDPRAPAGTICLTIRIVHSGRVRLQDFE--QIVSSSHDDTILMD 545
 DB 360 KTLRVWDF-----KMKRMKMTLNAHEHFTVSLDFHKTAPYVYTGSDQTVKWE 408
 RESULT 15
 Q9GL51 PRELIMINARY; PRT; 410 AA.
 AC Q9GL51; 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PLATELET-ACTIVATING FACTOR ACETYLAHYDROLASE IB-ALPHA SUBUNIT.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nagasaka T., Bouleday G., Coupel S., Coulon F., Tesson L.,
 RA Heslan J.-M., Souillhou J.-P., Charreau B.;
 RT "Cloning of porcine PAF-AH 1b-alpha cDNA and expression in endothelial
 RT cells";
 RL Submitted (NOV--2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF319658; AAG33867.1; -
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Hydrolase; Repeat; WD repeat.
 SQ SEQUENCE 410 AA; 46654 MW; A08DAFCDB88B2719 CRC64;

Query Match 11.6%; Score 351; DB 6; Length 410;
 Best Local Similarity 28.9%; Pred. No. 7.7e-21;
 Matches 86; Conservative 58; Mismatches 110; Indels 44; Gaps 8;

QY 280 IYSGLRDNTIKIMDNTLECKRILTGHTGSLVCLQYDE--RIYITGSSDSYRVWVDVNTG 337
 DB 123 IYASASDATIKVMYDETQDFERTLKGHTSYODISFDHSGKLKILASCSADMRTIKLMDQGF 182
 QY 338 EMLNTLIHCEAVLHRF--NNGMAYTCSKDRSIAVWDMASPTDITLRVVGHRAAVNV 395
 DB 183 ECIIRTMGHGDNHVSVAIMPNDHIVASARDKTIKWEVQTYCV---KTFGRHREVRM 239
 QY 396 V--DFDDKIYVASGDRITKIVWNTSTCEVFTLNGHKRGIAQLQYRDR----- 441
 DB 240 VRPNQDGLILASCSNDQYRVWVAVATKECKALREHENVYECISWAPESYSTSEATGS 299

QY 442 -----LVVSGSSDNTIRLMDIEGACLRVLEBGHELVRCIREDN--KRIVSGAYD 489
 Db 300 ETRKSGKPCPEFLLSGSRDKTIKMDVSTGCMITLVGHDMWVRGVLFFHSGKFILSCADD 359
 QY 490 GKIVMDLVAAIDPRAPAGTILCLTIVEHSGRVFRLOFDEF--QIVSSSHDPTILMD 545
 Db 360 KTRVWDY-----KNKRCMKTLNAAHEHFTSLDEHKTAPYVVTGSVDOTVAKWE 408

Search completed: May 8, 2002, 10:53:44
 Job time: 226 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:50:48 ; Search time 13.31 Seconds

(without alignments)
1567.414 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034

Sequence: 1 MDPAEAVLQEKALFKFNSS.....PAAQAEPPSPRTTYTISR 569

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3006	99.1	605	FWIA_HUMAN	Q9Y297 homo sapien
2	2597	85.6	518	TRCB_XENLA	Q91854 xenopus lae
3	2384.5	78.6	542	FWIB_HUMAN	Q9ubk1 homo sapien
4	1635.5	53.9	701	YSS1_CAEEL	Q09990 caenorhabdi
5	690	22.7	506	YAR1_SCHPO	Q09855 schizosacch
6	590.5	19.5	605	POF1_SCHPO	P87053 schizosacch
7	575	19.0	678	SCOB_EMENT	Q00659 emericella
8	545	18.0	640	MT30_YEAST	P39014 saccharomyc
9	531.5	17.5	650	SCO2_NEUCR	Q01277 neurospora
10	520	17.1	579	SE10_CAEEL	Q93794 caenorhabdi
11	477.5	15.7	684	CC4_CANAL	P53699 candida alb
12	455.5	15.0	1356	HEH1_PODAN	Q00808 podospira a
13	453	14.9	775	POPL_SCHPO	P87060 schizosacch
14	399	13.2	779	CC4_YEAST	P07834 saccharomyc
15	396.5	13.1	703	POB2_SCHPO	Q14170 schizosacch
16	392	12.9	732	KMBR_DICDI	P90648 dictyosteli
17	354	11.7	409	LISI_HUMAN	P43034 homo sapien
18	354	11.7	409	LISI_MOUSE	P43033 mus musculu
19	353	11.6	409	LISI_BOVIN	P43033 bos taurus
20	341	11.2	515	YCM2_YEAST	P25382 saccharomyc
21	337	11.1	1146	KMHA_DICDI	P42527 dictyosteli
22	334.5	11.0	422	FBW2_HUMAN	Q9ukt8 homo sapien
23	325.5	10.7	361	WDS_DROME	Q9v3f8 drosophila
24	324.5	10.7	362	WDR5_HUMAN	Q9u9p9 homo sapien
25	318	10.5	422	FBW2_MOUSE	Q60884 mus musculu
26	313.5	10.3	376	YKVA_CAEEL	Q17963 caenorhabdi
27	312.5	10.3	714	YU12_YEAST	P47025 saccharomyc
28	307.5	10.1	742	PKNA_THECU	P49695 thermomonas
29	307.5	10.1	1693	Y163_STY3	Q55663 synecocyst
30	306	10.1	704	T2D4_DROME	P49846 drosophila
31	299.5	9.9	1194	APAF_HUMAN	O14727 homo sapien
32	298.5	9.8	800	T2D4_HUMAN	O15542 homo sapien
33	292	9.6	659	YK16_YEAST	P36130 saccharomyc

34	290.5	9.6	614	1	TU11_SCHPO	Q09715 schizosacch
35	287.5	9.5	327	1	GBLP_BRANA	Q39336 brassica na
36	285.5	9.4	798	1	T2D4_YEAST	P38129 saccharomyc
37	283.5	9.3	327	1	GBLP_ARATH	O24456 arabidopsis
38	283	9.3	473	1	PRP5_SCHPO	O13615 schizosacch
39	283	9.3	713	1	TU11_YEAST	P16649 saccharomyc
40	281.5	9.3	682	1	TU11_KULIA	P56094 klyveromyc
41	280.5	9.2	444	1	NODE_EMENT	Q00664 emericella
42	280	9.2	318	1	GBLP_DROME	O18640 drosophila
43	277	9.1	586	1	TU12_SCHPO	Q9uug8 schizosacch
44	276.5	9.1	501	1	YH92_CAEEL	Q23256 caenorhabdi
45	273	9.0	325	1	GBLP_MEDSA	O24076 medicago sa

ALIGNMENTS

RESULT 1

FWIA_HUMAN STANDARD: PRT: 605 AA.

ID Q9Y297; Q9Y213;
AC 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP)
DE (E3RSIKAPPAE) (PIKAPPAALPHA-E3 RECEPTOR SUBUNIT).
GN FBWIA OR FBWIA OR BTRCP OR BTRC.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;

[1]
SEQUENCE FROM N.A. (ISOFORM 1).
MEDLINE=99075339; PubMed=9859996;
RA Yaron A., Hatzubai A., Davis M., Lavon I., Amit S., Manning A.M.,
RA Andersen J.S., Mann M., Mercurio F., Ben-Neriah Y.;
RT "Identification of the receptor component of the Ikapappaalpa-
ubiquitin ligase.";
RL Nature 396:590-594(1998).

[2]
SEQUENCE FROM N.A. (ISOFORM 2).
RP TISSUE=Lymphoid;
RX MEDLINE=98325370; PubMed=9660940;
RA Margottin F., Bour S.P., Durand H., Selig L., Benichou S., Richard V.,
RA Thomas D., Strebel K., Benarous R.;
RT "A novel human WD protein, h-beta TRCP, that interacts with HIV-1 Vpu
connects CD4 to the ER degradation pathway through an F-box motif.";
RL Mol. Cell 1:565-574(1998).

[3]
SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=20003060; PubMed=10531035;
RA Cenciarelli C., Chiau D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;
RT "Identification of a family of human F-box proteins.";
RL Curr. Biol. 9:1177-1179(1999).

[4]
CHARACTERIZATION.
RP MEDLINE=99145464; PubMed=9990852;
RA Winston J.T., Strack P., Beer-Romero P., Chu C.Y., Ellledge S.J.,
RA Harper J.W.;
RT "The SCF(beta-TRCP)-ubiquitin ligase complex associates specifically
with phosphorylated destruction motifs in I-kappa-B-alpha and
beta-catenin and stimulates I-kappa-B-alpha ubiquitination in vitro.";
RL Genes Dev. 13:270-283(1999).

-I- FUNCTION: BINDS SPECIFICALLY TO PHOSPHORYLATED IKB ALPHA
(PIKAPPAALPHA) AND PHOSPHORYLATED BETA-CATENIN AND PROMOTES THEIR
UBIQUITINATION AND DEGRADATION.
CC -I- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.
CC -I- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS, 1 (SHOWN HERE) AND 2; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -I- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF101784; AAD08702.1; -
 DR EMBL: Y14153; CAAT4572.1; -
 DR EMBL: AF129530; AAF04464.1; -
 DR MIM: 603482; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; WD repeat; Alternative splicing.
 FT DOMAIN 190 228 F-BOX.
 FT REPEAT 301 338 WD 1.
 FT REPEAT 341 378 WD 2.
 FT REPEAT 381 418 WD 3.
 FT REPEAT 424 461 WD 4.
 FT REPEAT 464 503 WD 5.
 FT REPEAT 505 541 WD 6.
 FT REPEAT 553 590 WD 7.
 FT REPEAT 590 605 MISSING (IN ISOFORM 2).
 SQ SEQUENCE 605 AA: 68866 MW: 46673B7E400FD37 CRC64;

Query Match 99.1%; Score 3006; DB 1; Length 605;
 Best Local Similarity 94.0%; Pred. No. 1,1e-219;
 Matches 569; Conservative 0; Mismatches 0; Indels 36; Gaps 1;

QY 1 MDPAAVILQEKALKFM-----NSSERDC 24
 DB 1 MDPAAVILQEKALKFMCMSPRLMIGCSSLADSMPLSLCINPGALITATONSSERDC 60
 QY 25 NNGEPPRKIPKNSLRQTYNSCARLCLNOETVCLASTAMTENCVAATKLANGTSSMTV 84
 DB 61 NNGEPPRKIPKNSLRQTYNSCARLCLNOETVCLASTAMTENCVAATKLANGTSSMTV 120
 QY 85 PRORLTSAYEKEKELCYKFFQWSESOVEVEHLISQMKHYOHGHINSLIKPMLQDF 144
 DB 121 PRORLTSAYEKEKELCYKFFQWSESOVEVEHLISQMKHYOHGHINSLIKPMLQDF 180
 QY 145 ITALPARGLDHIAENTLSYLDKSLCAAEVLCCKEMRYVTSQGLMKKLTIERVPTDSLIR 204
 DB 181 ITALPARGLDHIAENTLSYLDKSLCAAEVLCCKEMRYVTSQGLMKKLTIERVPTDSLIR 240
 QY 205 GLAERRGQGYLFKNKPPDGNAPNSFYRALYPTIIDIETIESNMWRCGRSLORICRS 264
 DB 241 GLAERRGQGYLFKNKPPDGNAPNSFYRALYPTIIDIETIESNMWRCGRSLORICRS 300
 QY 265 ETSKGVYCLQYDDOKIYSGLDNITIKTMDKNTLECKRLTGHITSVLCLOYDERVIITGS 324
 DB 301 ETSKGVYCLQYDDOKIYSGLDNITIKTMDKNTLECKRLTGHITSVLCLOYDERVIITGS 360
 QY 325 SDSYRVVDVNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSIAYWDNASPTDILRR 384
 DB 361 SDSYRVVDVNTGEMLNTLIHCEAVLHLRFNNGMAYTCSKDRSIAYWDNASPTDILRR 420
 QY 385 VLGVHRAAVNVVDFDDKYIVSASDRTIKWNISTCFEVRTLNGHKGIACTQYRDLVV 444
 DB 421 VLGVHRAAVNVVDFDDKYIVSASDRTIKWNISTCFEVRTLNGHKGIACTQYRDLVV 480

QY 445 SGSSDNTLRIMDIECGACLRVLEBGEELVRCIRPDNRKRIYSGAYDKIKWVDVYALDPR 504
 DB 481 SGSSDNTLRIMDIECGACLRVLEBGEELVRCIRPDNRKRIYSGAYDKIKWVDVYALDPR 540
 QY 505 APAGTLICRLTVESGVRFLQFDEFOIVSSSHDITLIMDFLNDPAAQAAPPSPRSRTY 564
 DB 541 APAGTLICRLTVESGVRFLQFDEFOIVSSSHDITLIMDFLNDPAAQAAPPSPRSRTY 600
 QY 565 TYISR 569
 DB 601 TYISR 605

RESULT 2
 TRCB_XENIA STANDARD; PRT; 518 AA.
 ID TRCB_XENIA
 AC Q91854; P70037; P70038;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE BETA-TRCP (BETA-TRANSDUCIN REPEAT-CONTAINING PROTEIN).
 GN FBXW1 OR BTRCP.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9330289; PubMed=8393141;
 RA Spevak W., Keiper B.D., Stratowa C., Castanon M.J.;
 RT "Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in
 RT anaphase are rescued by Xenopus cDNAs encoding N-ras or a protein
 RT with beta-transducin repeats";
 RT Mol. Cell. Biol. 13:4953-4966(1993).
 RL [2]
 RP SEQUENCE OF 302-518 FROM N.A.
 RX MEDLINE=97109804; PubMed=8952061;
 RA Hudson J.W., Alarcon V.B., Elinson R.P.;
 RT "Identification of new localized RNAs in the Xenopus oocyte by
 RT differential display PCR";
 RT Dev. Genet. 19:190-198(1996).
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PHOSPHORYLATED
 CC PROTEINS AND PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CUL1IN-F-BOX) PROTEIN LIGASE COMPLEX
 CC (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: PRESENT IN FULLY GROWN AND PROGESTERONE-
 CC MATURED OOCYTES. THE LEVEL CHANGE VERY LITTLE EVEN AFTER ZYGOTIC
 CC GENE TRANSCRIPTION BEGINS FOLLOWING THE MIDBLASTULA TRANSITION. DO
 CC NOT INCREASE IN ABUNDANCE IN THE GASTRULA, NEURULA, TAILBUD, OR
 CC TADPOLE EMBRYO.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -1- SIMILARITY: STRONG, TO C.ELEGANS K10B2.1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M98268; AAA02810.1; -
 DR EMBL: U63921; AAA9671.1; -
 DR EMBL: U63922; AAA9672.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.

DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS00678; WD_REPEATS_1; 6.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; WD repeat.
 FT DOMAIN 119 157 F-BOX.
 FT REPEAT 230 258 WD 1.
 FT REPEAT 270 298 WD 2.
 FT REPEAT 310 338 WD 3.
 FT REPEAT 353 381 WD 4.
 FT REPEAT 393 421 WD 5.
 FT REPEAT 433 461 WD 6.
 FT REPEAT 482 510 WD 7.
 FT REPEAT 502 530 GEM -> EFR (IN REF. 2).
 FT CONFLICT 302 304 GLA -> AAH (IN REF. 2).
 FT CONFLICT 516 518
 FT SEQUENCE 518 AA; 59507 MW; 2A52EC19028127E3 CRC64;
 Query Match 85.6%; Score 2597; DB 1; Length 518;
 Best Local Similarity 91.6%; Pred. No. 7.2e-189;
 Matches 488; Conservative 7; Mismatches 8; Indels 30; Gaps 1;
 QY 18 SEERECNNGEPPRIPEKNSLRQTVNSCARLCLNOETVCLASTAMTENCVAKTLAN 77
 DB 13 ASERECNDEPPKITEKNTLRO-----TKLAN 42
 QY 78 GSSMIVPQKRLASYEKEKELCVKPEWSESDVEFEHLISOMCHYGHINSYLK 137
 DB 43 GSSMIVPQKRLASYEKEKELCVKPEWSESDVEFEHLISOMCHYGHINSYLK 102
 QY 138 PMLQDFITAPARGLDIAENITSLDAKSLCAAEVLCKEYRTSDGMKLIERAY 197
 DB 103 PMLQDFITAPARGLDIAENITSLDAKSLCAAEVLCKEYRTSDGMKLIERAY 162
 QY 198 RFDLSLWGLAERGMQYLFKNKPPGNAPNSFYALPKITIDETTESNMGRGHS 257
 DB 163 RFDLSLWGLAERGMQYLFKNKPPGNAPNSFYALPKITIDETTESNMGRGHS 222
 QY 258 QRIHCRSETSKGYVCLQYDQKIVSGRLDNTIKIMDKNTLECKRIITGHTGSLCLQYDE 317
 DB 223 QRIHCRSETSKGYVCLQYDQKIVSGRLDNTIKIMDKNTLECKRIITGHTGSLCLQYDE 282
 QY 318 RVIITSSSSTYRVMDVNTGEMLNTLIHHCVAVLHFRNNGMMVTCSSKRSIAVMDMASP 377
 DB 283 RVIITSSSSTYRVMDVNTGEMLNTLIHHCVAVLHFRNNGMMVTCSSKRSIAVMDMASP 342
 QY 378 TDTTLRRVIVGHRAAVNVDPDDKIVYSAGPRTIKVMTSTCEPRTINGHKGRIACIQ 437
 DB 343 TDTTLRRVIVGHRAAVNVDPDDKIVYSAGPRTIKVMTSTCEPRTINGHKGRIACIQ 402
 QY 438 YRDLVYSSGSDNTIRLMDIEGACLVLEHGEHLVRCIRFNNKRIVSGAYDGKIKVMDL 497
 DB 403 YRDLVYSSGSDNTIRLMDIEGACLVLEHGEHLVRCIRFNNKRIVSGAYDGKIKVMDL 462
 QY 498 VALADPRAGTICLTIVHSGRVRLQFDEFQIVSSHDTILIMDLNDP 550
 DB 463 VALADPRAGTICLTIVHSGRVRLQFDEFQIVSSHDTILIMDLNDP 515
 RESULT 3
 FW1B_HUMAN STANDARD: PRT: 542 AA.
 AC Q9UKB1; Q9Y4C6; Q9P2S8; Q9P2S9;
 DT 20-AUG-2001 (Rel. 40; Created)
 DT 20-AUG-2001 (Rel. 40; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE F-BOX/WD-REPEAT PROTEIN 1B (F-BOX AND WD-REPEATS PROTEIN BETA-TRCP2).
 GN FBXW1B OR FBXW1B OR BTRCP2 OR KIA0696.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE-20003060; PubMed-10531035;
 RA Chantrelle C., Chitaur D.S., Guardavaccaro D., Parks W., Vidal M.,
 RA Pagano M.;
 RT "Identification of a family of human F-box proteins.";
 RL Curr. Biol. 9:1177-1179(1999).
 RN [12]
 RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 RC TISSUE=Fetal lung;
 RX MEDLINE-20160458; PubMed-10694485;
 RA Koike J., Sagara N., Kirikoshi H., Takagi A., Miwa T., Hirai M.,
 RA Katoh M.;
 RT "Molecular cloning and genomic structure of the betatrop2 gene on
 RT chromosome 5q35.1.";
 RL Biochem. Biophys. Res. Commun. 269:103-109(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE-98403880; PubMed-9734811;
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998)
 CC -1- FUNCTION: PROBABLY RECOGNIZE AND BINDS TO SOME PROTEINS AND
 CC PROMOTES THEIR UBIQUITINATION AND DEGRADATION.
 CC -1- SUBUNIT: PART OF A SCF (SKP1-CULLIN-F-BOX) PROTEIN LIGASE COMPLEX.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A, B AND C (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF176022; AAF04528.1; -;
 DR EMBL: AB033279; BAA92329.1; -;
 DR EMBL: AB033280; BAA92330.1; -;
 DR EMBL: AB033281; BAA92331.1; -;
 DR EMBL: AB014596; BAA31671.1; ALT_INIT.
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS: PR00320; GPROTEINBRPT.
 DR SMART: SM00256; FBOX; 1.
 DR SMART: SM00320; WD40; 7.
 DR PROSITE: PS50181; FBOX; 1.
 DR PROSITE: PS50082; WD_REPEATS_1; 5.
 DR PROSITE: PS50082; WD_REPEATS_2; 7.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 KW Ubiquitin conjugation; Repeat; Alternative splicing.
 FT DOMAIN 129 167 F-BOX.
 FT REPEAT 129 167 WD 1.
 FT REPEAT 238 275 WD 2.
 FT REPEAT 278 315 WD 3.
 FT REPEAT 318 355 WD 4.
 FT REPEAT 361 398 WD 5.
 FT REPEAT 401 440 WD 6.
 FT REPEAT 442 478 WD 7.
 FT REPEAT 490 527 WD 8.
 FT REPEAT 527 564 WD 9.
 FT VARSPLIC 16 49 MISSING (IN ISOFORM A).
 FT VARSPLIC 16 48 CSVRSLIMICANIVESMCALSCSMPVRL -> NTSV
 FT SEQUENCE 542 AA; 62090 MW; 7CD40087EFAA5C8A CRC64;
 SQ

```

or send an email to license@slsb.cb).
-----
CC      EMBL: U028730; AAA68258.1; -
DR      WormRep; K10B2.1; CE02008.
DR      InterPro; IPR001810; F-box.
DR      InterPro; IPR001680; WD40.
DR      Pfam; PF00646; F-box; 1.
DR      Pfam; PF00400; WD40; 7.
DR      PRINTS; PR00320; GPROTEINBRPT.
DR      SMART; SM00256; FBOX; 1.
DR      SMART; SM00320; WD40; 7.
DR      PROSITE; PS0181; FBOX; 1.
DR      PROSITE; PS00678; WD_REPEATS_1; 5.
DR      PROSITE; PS50082; WD_REPEATS_2; 7.
DR      PROSITE; PS50294; WD_REPEAT_REGION; 1.
KW      Hypothetical protein; Repeat; WD repeat.
FT      REPEAT                256   284       WD 1.
FT      REPEAT                296   324       WD 3.
FT      REPEAT                336   364       WD 2.
FT      REPEAT                379   407       WD 4.
FT      REPEAT                419   447       WD 5.
FT      REPEAT                459   487       WD 6.
FT      REPEAT                508   536       WD 7.
FT      DOMAIN                 606   615       POLY-ALA.
SQ
SEQUENCE    701 AA:  80320 MW:  69FA0B0F83270E3 CRC64:

Query Match          53.9%; Score 1635.5; DB 1; Length 701;
Best Local Similarity 57.7%; Pred. No. 4,7e-116;
Matches 328; Conservative 69; Mismatches 116; Indels 55; Gaps 8

QY      31 KRITPEKNSLRQTYSCARCLNQEYVCLASTAMKTENCVAKTKLAN-----GTS 80
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      2 RFRREGKRALKGRRARDGSGIAQLTVCVST----IERCF--TAVSNPIFFLFSEFSYF 55

QY      81 SNAIVEKOR-----KLASAYEKELCKVKFEQMGSSDOVEFEYHILISOMCHOHGHINSX 135
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      56 SLPLFSRNTQIPLSRSSFSSEYVL-----KWEHEQIDFMDKYVHRISHIQLGKVDNF 110

QY      136 LKPMLQDEFTALPARGDITHAENTLSYLDAKSLCAELVCKEWRYVSDGMIMKKLIER 195
        ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      111 IRPMLQGFIFSMPLA-----HLVELILFNVSDSLKSCEVESTSWRCALARGGHMKKLEK 166

QY      166 MYRTSIRNGLAERGMGQYL-----PKNNPPDGNAPPNSFYRL 235
        ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      167 NVRSDSLMWGLSEKRQMDFLNISKDMSVRICERFNYDVNIKRDQLDLIMHFYSKL 226

QY      236 YPKIIODIEETIESNMCRGRHSIORICHSETSKGYCLOQYDDOKIVSGLRDTIKIPDN 295
        ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      227 YPKIIRDHNINDNNKRGYKMTIRINCQSENKGYCCLOQYDDDKIVSGLRDTIKIWDRK 286

QY      296 TLECRIILTGHGTGVLCQYDERVIITGSSDSITYRWMDVTGEMLTLIHCBAVLHREF 355
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      287 DYSCRSRIISGHGTGYLCQYDNRVITIGSSSPATVWMDVEGEICIKTLIHCEAVLHREF 346

QY      386 NGGMWWTOSKPSRTAWMVMASPTDITLRVLYGHHAAVNVDPODKITYSAGDPTIKYW 415
        ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      347 ANGLIWITSKDPSIAWVMWVSPRDIITIRLVLGHPAAVANVDEDTRYIVSAGDPTIKYW 406

QY      416 NTSTICEFVRTLNHGHRGTACIQYRDLRVLVSGSSDNTIILAMDIECGACIRVLEGHEBELRC 475
        ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      407 SMDTLEFVRTLAGHHRGTACIQYRGLRVLVSSSDPTIKLMDIHSGVCLRVLEGHEBELYRC 466

QY      476 TRFDNKRIYSAGDYCKIKYMDLVLAALDPRAPAGTLCRLTVEHSGVRVRLFDFEQIVSS 535
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      467 TRFDEKRIYSAGDYCKIKYMDLVLAALDPRALSSEICLSIOHTGVRVRLFDFEQIVSS 526

QY      536 SHDDTIILMDFLNPAPAAQAEPPRSPT 563
        |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      527 SHDDTIILMDFLDAP-----PSGLPST 549

```


YAF1_SCHPO STANDARD: PRT: 506 AA.

AC 009855: 09P7V1: (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHEICAL 58.3 KDA TRP-ASP REPEATS CONTAINING PROTEIN C2956.01 IN CHROMOSOME 1.

GN SPAC2956.01 OR SPAC30.05.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

OC NCB1_TaxID=4896;

OX NCB1_TaxID=4896;

RN NCB1_TaxID=4896;

RP STRAIN-972;

RC McDougall R.C., Rajandream M.A., Barrell B.G., Brown S., Murphy L., Jones L., McNeil A., Harris D.,

RA Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

RL [2]

RN SEQUENCE OF 240-506 FROM N.A.

CC STRAIN-972:

CC Jones L., Murphy L., McNeil A., Simpson I., Harris D., Barrell B.G., Rajandream M.A., Walsh S.V.,

RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: A1136538; CAB66464.1; -

DR EMBL: 266525; CA91423.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

DR PRINTS: PR00320; GPROTEINBPT.

DR SMART: SM00256; FBOX; 1.

DR SMART: SM00320; WD40; 7.

DR PROSITE: PS00181; FBOX; 1.

DR PROSITE: PS00678; WD_REPEATS_1; 3.

DR PROSITE: PS50082; WD_REPEATS_2; 6.

DR PROSITE: PS50294; WD_REPEATS_REGION; 1.

KW Hypothetical protein; Repeat; WD repeat.

KW DOMAIN

FT REPEAT 219 256 WD 1.

FT REPEAT 259 298 WD 2.

FT REPEAT 301 338 WD 3.

FT REPEAT 345 386 WD 4.

FT REPEAT 388 426 WD 5.

FT REPEAT 427 464 WD 6.

FT REPEAT 468 505 WD 7.

SO SEQUENCE 506 AA; 58257 MW; CER34D4EFPBC2E10 CRC64;

Query Match 22.7%; Score 690; DB 1; Length 506;

Best local similarity 30.4%; Pred. No. 8.3e-45;

Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 ENCVAR-----TKANGTSSNIYPKOKKLSAYKEKELCVKFEQMSSEDQVEFVHLIS 122

DB 8 KNVASVSDLTSCDSTSPVCLNPLS-----HENNRIDLRLDLA 50

QY 123 OMCHYHGHTINSYLRKMLDQFTALPARGLDHIAENILSYLDKSLCAAEIIVCKEMRYV 182

DB 51 SLKREGVAVYNNHVRSLFTFDTEVRP-----EVSILRVSYLDQDLCKCKLMSKRMKRL 106

QY 183 TSDGLMKKLI-----ERMVFTDSLWRG-----LAERRGNG----- 213

DB 107 LEDGIMKALYMKQGWPFVNEVNLNEFEAMRTHKFPQPRENFKLQOONITPGTFLPQ 166

QY 214 QYLEKKNPPGNNAPNPFYALPKIQTIDLETIESMWRCSRSLQRIHCS----- 264

DB 167 QEIF-----DSNGRPLLNWSLY--KEHAHDSNNRHRFLVSTFNPNSTIRPADQF 217

QY 265 -ETSKGVCLQYDPOKIVSGLRDNTIKIMDKNTLECKRIITGHTGSLQDYDER--VII 321

DB 218 RATIDSVYCYQYDEIVSGSKDRTVMDVNSFFILKLYHSGSVLCLDFCRRLRLV 277

QY 322 TGSSDSTVRWVDVNTGEMLNTLHCEAVLHLRFNNGMAYTCKSDSIAYV--DMSPTD 379

DB 278 SGSSDSTIIIDWQNRRLPKYFGHTDNLGVVSENYIISSSDRTARWRADATSPAE 337

QY 380 ITRRLVGHRAAVNVYDFDCK--YIVASGDRIRIKWNTSTCEFTVTLNGHKGIACIQ 437

DB 338 ACM-HVLRGHLASVNSVQYSKGTGLIVASSDRILRTWDLTTGICITIIAHORGICAQ 396

QY 438 YRDLVYSGSSDNTIRLMDIEGACLVLEGHELVRCIRFNKRIYSGAYDKIKYMDL 497

DB 397 YNGKFIVSGSSDLIRIFEASSGKLRLYLGHEDLIRTFVFNDEKIVSGYDGTVRIMN- 455

QY 498 VALDPRAPACTLCRLRIVER-----SGRVFLQDFDEQIVSSSHDITLLWDF 546

DB 456 -----FNTGECHVLHNSRNSRVFGLQDFHRIITACTHSSEILWNF 497

RESULT 6

POFI_SCHPO STANDARD: PRT: 605 AA.

AC P87053;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE F-BOX/WD-REPEAT PROTEIN POFL.

GN POFL OR SPAC57A10.05.C.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetales; Schizosaccharomycetaceae; Schizosaccharomycetes.

OC NCB1_TaxID=4896;

RN NCB1_TaxID=4896;

RP STRAIN-972;

RC Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;

RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

RN NCB1_TaxID=4896;

RP STRAIN-972;

RC Badcock K., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;

RA Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.

CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC -----

CC EMBL: AB032410; BA84528.1; -

DR EMBL: 294864; CAB08168.1; -

DR InterPro: IPR001810; F-box.

DR InterPro: IPR001680; WD40.

DR Pfam: PF00646; F-box; 1.

DR Pfam: PF00400; WD40; 7.

PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 2.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 107 153 F-BOX.
 FT REPEAT 271 299 WD 1.
 FT REPEAT 311 339 WD 2.
 FT REPEAT 350 379 WD 3.
 FT REPEAT 390 420 WD 4.
 FT REPEAT 432 460 WD 5.
 FT REPEAT 472 500 WD 6.
 FT REPEAT 510 538 WD 7.
 SQ SEQUENCE 605 AA; 67110 MW; 7118C9379EC5C1F0 CRC64;

Query Match 19.5%; Score 590.5; DB 1; Length 605;
 Best Local Similarity 29.0%; Pred. No. 3,5e-37;
 Matches 156; Conservative 86; Mismatches 189; Indels 107; Gaps 12;

QY 80 SSMATPKQKSLASYTEKEKELCVKFEQWSE---SDQVEVHLLISOMCHYQHCHINSYL 136
 DB 46 SSM---HNELSGLESEKORVEAWAFAFSESCERKIALQGLINCSLSLFSASTL 101
 QY 137 KPMLODFITLAPAGLDHIAENILSYDAKSLCAELVCKEYRVTSQGLMKLIFRM 196
 DB 102 DSLVRLDLSLPV---EISFRILSFIDANSLCAOAGVSKHMKELADDVIMHRCBOH 157
 QY 197 VRTDLMKGLAERRGCGOYLEFN----- 219
 DB 158 INRK-----CEKCGWGLPLEERNTLYAAKASIQKRYERLTFRGVDAHESSPYKAKLD 211
 QY 220 -----KPPGNAPPNF-----YRALPKTIIDLETESMWRGRHSLO 258
 DB 212 DPTSSNETTSSVAKPPSPNSDSKFLPCKTRPKMEVYAEKRCR---VECNWRGR--- 263
 QY 259 RINCR---SETSKGVYCLQYDDQKIYSGRLDNTIKIMDKTLLCKRILTGTSVCLQ 314
 DB 264 ---CROVVLGSHSDGVMCLQVLRNIIASGSYDATIRLMNLTPOVALLEHSSGVCLQ 320
 QY 315 YDERIITIGSSDYRYVWDVNTGEMLNTLHHCEAVLHLPNNGMAYTCSKDRSIAYWDM 374
 DB 321 FDOCKLIGSMDKTRIRNWRYSSECISILHGTDSVCLTDFDSTLIGSADCTVAKLHF 380
 QY 375 ASPDITLRLVYGHRAAVNVDF--DDKYTVASGDRITKVMWMTSCFEFRTLNGHKG 432
 DB 381 SGGKRITLR---GHTGVNSVRIIRRGVLSGSDSTIKIWSLENTCLHTSAHGP 436
 QY 433 IACLOYRDLVYSGSSDNTIRLMDIEGACLVLEGHEELVRCIRFDRKRVSGAYDGI 492
 DB 437 VQSLALADSRLEFSCSLDGTIKOMDIEKKKCVHTLFGHEIGVEIAADHLRLISAHGAV 496
 QY 493 KVMIDVALDPRAGTICLTLEHSGRVRRLQDFEQIYSSSHDITILMDLNP 550
 DB 497 KVMEEACE-----CVHTLKNHSEPTVALGDEEVSGSDCKIYIWLFPNAP 543

RESULT 7
 SCOB_EMENT
 ID SCOB_EMENT STANDARD; PRT; 678 AA.
 AC 000659;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SULFUR METABOLITE REPRESSION CONTROL PROTEIN.
 GN SCOB OR MAPPI.
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryotes; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_Taxid=5072;

FN [1]
 RE SEQUENCE FROM N.A.
 RA Notoif R.
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: REGULATORY PROTEIN THAT CONTROL SULFUR METABOLITE
 CC REPRESSION.
 CC - SIMILARITY: CONTAINS 8 WD_REPEATS (TFR-ASP DOMAINS).
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC - SIMILARITY: BELONGS TO THE MET30/SCOB/SCON-2 FAMILY OF WD-REPEAT
 CC PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: U21220; AAC15905.1; -
 DR InterPro: IPR001810; F-box.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00646; F-box; 1.
 DR Pfam: PF00400; WD40; 7.
 DR PRINTS; PR00320; GPROTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 7.
 DR PROSITE; PS50181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS50082; WD_REPEATS_2; 7.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Transcription regulation; Repeat; WD repeat.
 FT DOMAIN 178 224 F-BOX.
 FT REPEAT 347 375 WD 1.
 FT REPEAT 387 415 WD 2.
 FT REPEAT 427 455 WD 3.
 FT REPEAT 466 496 WD 4.
 FT REPEAT 508 543 WD 5.
 FT REPEAT 553 595 WD 6.
 FT REPEAT 607 635 WD 7.
 FT REPEAT 647 675 WD 8.
 SQ SEQUENCE 678 AA; 76070 MW; D840D452E37B4C53 CRC64;

Query Match 19.0%; Score 575; DB 1; Length 678;
 Best Local Similarity 26.9%; Pred. No. 6e-36;
 Matches 149; Conservative 83; Mismatches 168; Indels 154; Gaps 14;

QY 117 VEHLISOMCHYOHGHTNSYLKPMLODFITLAPAGLDHIAENILSYDAKSLCAELVC 176
 DB 153 LOGIMAQCEFPOLSYISATVRDLIRIDETALP---PEIAFKIICLYDITSLCKASQVS 208
 QY 177 KEMRYTSGMLMKKLIEMVVRTDSLMRGLAERRGGOYLEFNK----- 220
 DB 209 KGRALADDDVYHNRKCEGHIRK-----CKKCGWGLPLEDRKRLRSKREIELRATW 262
 QY 221 -----PPDGA--PP-----NSFYRALY---P 237
 DB 263 DRGVVPRSPDASAESPPSGKRKLEDDVAVYKRRCSLSGSDAGVDKSDFFKTRYRPMK 322
 QY 238 KIIDLETESMWRGRHSLOIRHCRSEK---GVYCLQYDDQKIYSGRLDNTIKIWD 293
 DB 323 EYKDKFVKGYMKYGR-----CSIKTFKGTNGVMCLQEDNIIATGSDYDTIKIWD 375
 QY 294 KNTLECKRILTGTSVCLQYDERVIITIGSSDYRYVWDVNTGEMLNTLHHCEAVLHL 353
 DB 376 TEGEELFRLRHSIRGLQDFDQKILSGSMDRIKVMNWRYSSECISITGYGHGAVIGL 435
 QY 354 RNNNGMAYTCSKDRSIAYWDMASPTDITLRLVYGHRAAVNV--VDFDKYIVASGDR 411
 DB 436 HFDASILASGVADKTVKIMNFKSTFSLR---GHTDMVNAVRVYSSRTVFASDQCT 491
 QY 412 IKVMNTSTCEFRITLNGH-----KRIIAC----- 435

```

Db 492 VRLMDLDTKTCIRTFHGHVGVQVAVPLPRFEFEHHAECENDLSTTSGDANPPSIOA 551
OY 436 -----LOYRD-----RLVYSSGSSDNTIRLMDIEGACLRVLEGBEL 472
Db 552 SNGLEPNAVYSSGSAFSGSPDNGRAAPPYRYWTSALDSTIRLMTETTGTCRLTFEFGHLEB 611
OY 473 VRCIRFDPNRIYSGAVDGRKIKYMDLVAAALDPPAPAGTICLRTLVHSGRVRFLQDFEQI 532
Db 612 VVALGADTRIRIYSGAEDNRKIKMD-----PRTGKCEIRFTGHSRVCIGIGDSRF 662
OY 533 VSSSHDDTLLIMDF 546
Db 663 ATGSEDCVVRMYSF 676

RESULT 8
M30_YEAST
ID M30_YEAST STANDARD; PRT; 640 AA.
AC P39014;
DT 01-FEB-1995 (rel. 31, Created)
DT 01-FEB-1995 (rel. 31, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE MET30 PROTEIN
GN MET30 OR YIL046W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=42180-1A;
RA MEDLINE=96069360; PubMed=8524217;
RA Thomas D., Kuras L., Barbey R., Charest H., Blaiseau P.L.,
RA Surdin-Kerjan Y.;
RT "Met30p, a yeast transcriptional inhibitor that responds to S-
RT adenosylmethionine, is an essential protein with WD40 repeats."
RN [2]
RL Mol. Cell. Biol. 15:6526-6534(1995).
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrill B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
RA Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA Welsh S.V., Whitehead S.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR AMINO ACIDS BIOSYNTHESIS
CC GENES EXPRESSION.
CC -1- SUBUNIT: SEEMS TO INTERACT WITH MET4.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONB/SCON-2 FAMILY OF WD-REPEAT
CC PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: Z46861; CAAB6905.1; -
CC SGI: L26505; AAA96717.1; -
CC SGD: S0001308; MET30.
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00646; F-box; 1.
CC Pfam: PF00400; WD40; 6.
CC PRINTS: PR00320; GPROTEINBAPT.
CC SMART: SM00256; FBOX; 1.

```

```

DR SMART: SM00320; WD40; 6.
DR PROSITE: PS50181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 4.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Methionine biosynthesis;
KW Cysteine biosynthesis; Repeat; WD repeat.
FT DOMAIN 181 227
FT REPEAT 300 328 WD 1.
FT REPEAT 340 368 WD 2.
FT REPEAT 380 408 WD 3.
FT REPEAT 419 449 WD 4.
FT REPEAT 461 499 WD 5.
FT REPEAT 509 538 WD 6.
FT REPEAT 550 578 WD 7.
FT REPEAT 607 635 WD 8 (POTENTIAL).
FT REPEAT 61 M -> I (IN REF. 1).
FT CONFLICT 61
SQ SEQUENCE 640 AA; 72835 MW; 5135D4BCA2E1EB97 CRC64;

Query Match 18.08; Score 545; DB 1; Length 640;
Best Local Similarity 28.98; Pred. No. 1e-33;
Matches 167; Conservative 85; Mismatches 202; Indels 124; Gaps 17;

OY 34 IPEKNSLRQTYNSCARLCLN-QETVCLASTAMKTENCYAK-TRLANGTSSMIVPKORRLS 91
Db 86 LPEVFTKFCYTRHNDIPFSPYTHACRYKQDLKRTQELINANAKLPLOBQSDIHIIISKYS 145
OY 92 ASYEKEKELCVKYEQWSESDQVEFEHLISOMCHYOHGHSYLYKPMLODFITALPAR 151
Db 146 NSNDKIRKL-----ILDGLSTSCFPOLSTYSSIVTHMIKIDFISILP-- 188
OY 152 GLDHAENILSLDAKSLCAELNCKEWEYRTSGMIMMKKLIEMVRRLDSLMRGLAERR- 210
Db 189 --QELSLKILSLDQSLCNATVCRKWKQKLADDRVYHACEOH-----DRKC 236
OY 211 ---GNGOYLKFKK-----PPDGNAPNSFYKALPKIYDIETIESNMRCGRSLQ 258
Db 237 PNCGGLPLHMKRRARIRIOONSTGSSNMADIQTQTRPKVYIREBFKESNMRRG----- 291
OY 259 RIHCSETSK---GVCLQYDDQKIVSGLRDNITKIDKNTLECKRLITGHTSVLCLQ 314
Db 292 --HCRIOEFKGMHDVLLLOEFNYRLFLFGSYDSTIGIMDLFTGKLIRLSGHSOGVKTLY 349
OY 315 YVERIITIGSSSYRVMDVNTGEMLNLHHCFAVLHFRNNGMAMVCSKDRSLAVDM 374
Db 350 FDDRKLITGSLDKTIRVMNYITGECISTYRGHSVLSVDSYQKIVYSGADKYKVMHV 409
OY 375 ASPTITLRLVLVGHRAAVNVVDFDKYT--VSASGDRITKVMNTSGEYVRLNGH-- 429
Db 410 ESRCTYTLR-----GHTENVNCKYKLPKFSFCSCSDDTTIRMMDIRNTSCLKFRGHVQ 465
OY 430 KRGIACLOYRD--RLVSSGS----- 448
Db 466 VOKIIPITKIDVENLATNTSDSSSPDDPTWTDGADSDPPSNQETVLDENITPYPTL 525
OY 449 -----DNTIRLMDIEGACLRVLEGBHELYRCIRFDRNRIYSGAVDGRKIKYMDLVAAALDP 503
Db 526 LSCGLDNTIKLMDVTKGKIRTFQFHVBEWDIADNRIISGSHDSIKYMDLQSG--- 582
OY 504 RAPAGTICLRTLVHSGRVRFLQDFEQIVSSSDTI 541
Db 583 -----KCMHTF---NGR--RLQRETQHTQTSGLDKV 609

RESULT 9
SC02_NEUCR
ID SC02_NEUCR STANDARD; PRT; 650 AA.
AC 001277;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE SULFUR CONTROLLER-2 (SCON2).

```

```

GN SC0N-2.
OC Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OC NCBI_TaxId=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=74-OR23-1A;
RX MEDLINE=95241499; PubMed=7724564;
RA Kumar A.; Paletta J.V.;
RT "The sulfur controller-2 negative regulatory gene of Neurospora
  crassa encodes a protein with beta-transducin repeats.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:3343-3347(1995).
CC -1- FUNCTION: NEGATIVELY REGULATES SULFUR STRUCTURAL GENE EXPRESSION.
CC -1- INDUCTION: EXPRESSED ONLY UNDER LOW-SULFUR CONDITIONS.
CC -1- SIMILARITY: CONTAINS 8 WD REPEATS (TRP-ASP DOMAINS).
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE MET30/SCONE/SCON-2 FAMILY OF WD-REPEAT
  PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U17251; AAA68968.1;
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS0181; FBOX; 1.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS00682; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Transcription regulation; Repeat; WD repeat.
FT DOMAIN 124 170
FT REPEAT 292 320 WD 1.
FT REPEAT 332 360 WD 2.
FT REPEAT 372 400 WD 3.
FT REPEAT 411 441 WD 4.
FT REPEAT 453 488 WD 5.
FT REPEAT 528 564 WD 6.
FT REPEAT 576 604 WD 7.
FT REPEAT 616 644 WD 8.
SQ SEQUENCE 650 AA; 72189 MW; 7473859C99F1B028 CRC64;

```

Query Match 17.5%; Score 531.5; DB 1; Length 650;

Best Local Similarity 25.0%; Pred. No. 1,1e-32;

Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

```

QY 117 VEHLISQMHYQHGHINSYKPKLQRFDTALPARGLDHTAENTLSTYDAKSLCAAEIYC 176
DB 99 LQGLISOLCPDPSFVSEVNEALKIDFISALPV---ELAQVYLCLDPLVSLTKAAQVS 154
QY 177 KEWRYVTSDDMLMKLIERVVRDLSLIRGLAEKRGQYLFRRKP-----P 222
DB 155 QRRMTIADSLAVVVRMCEQHVNKR-----CTKCGWGLPILERRKLLNVTROQLAKGGP 208
QY 223 DQNA-----PNSFYRA 234
DB 209 QGRVTELDSDSDSVNOHGKRPAAEAEDPPIKKRQCMMAAEASKAVTQKTSWMA 268
QY 235 LYKRIIDITFIESNMRCGRHSIQRIHCHSETSKGYCQIYDDOKIVSGIRNDTIRIMDK 294
DB 269 VY-----RDRQVSYNMKNSRYKLTVL---KGHENGVTCLQDLDNIIATGSDYPTIKIWI 321

```

```

QY 295 NILECKRIILFHTGVSVCLODYDERVIITGSSDSTVRWVDVNTGEMLNTLLIHCEAVLHR 354
DB 322 ETEECIRTLWGHAGIRALQFDDSKLISGSLDHTIKYWMHTGECISTFAAHDVSIVH 381
QY 355 FNNGMVTCSDKRSIAVMDASPTDILRVLVGHAAVNV--VDPDDKIYISASGDRIT 412
DB 382 FDGHLIASGSSDKTVKIFPDSKETCYLCK---GHSQWVNSTHVDIKSRVFSASDDTYI 437
QY 413 KVMNTSTCEVFRTLNGHKRGI-----ACLQYRDLVVS----- 445
DB 438 KLMDLDRQYIFRYEGHGHVQOVLILPRYEPDEVLNGASQDNDAMSVSQSGSGSPS 497
QY 446 -----GSSD 449
DB 498 MSHAQIERAGSPGSHSSSHNLPPSLPSGDEDEVRLYGSAPVADESRLPPRYFMTGGLD 557
QY 450 NTRILMDIEGACLRVLEGHLEELVRCIRPDNKRIVGAYDGKIKVMDLVAALDPAPAGT 509
DB 558 STMRLWDSATGRCLRTFLFGLBECVWSLADGTIRVIGANDGVKTYE-----PRSG 608
QY 510 LCLRTLVHSGRVRLQFDPFOIVSSHDDTILIMDF 546
DB 609 KCDATYTGHCVPYTCVGLSDSIMASGSEDGTTIRLSHF 645
RESULT 10
SEIO_CABEEL
AC 093794;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SEL-10 PROTEIN
GN SEL-10 OR F55B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Sims M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RC STRAIN-BRISTOL N2;
RA Jones S.J.M.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z79757; CAB02129.1;
DR WormRep: F55B12.3; CE16120.
DR InterPro: IPR001810; F-box.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00646; F-box; 1.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR SMART: SM00256; FBOX; 1.
DR SMART: SM00320; WD40; 7.
DR PROSITE: PS00678; FBOX; 1.
DR PROSITE: PS00682; WD_REPEATS_1; 5.
DR PROSITE: PS50082; WD_REPEATS_2; 7.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.

```


OY 506 PAGTLCITLVHSGRVERLQD-----EFOIYSSSHDITL-IMDPLNDPAAQ 553
 DB 630 EKGTAFRLDLSVTGGIMQVRFYKCAVAQVITNEDECEFIIDLF-SQPLIQ 684

RESULT 12

HET1_PODAN STANDARD; PRT: 1356 AA.
 ID HET1_PODAN

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE VEGETATIVE INCOMPATIBILITY PROTEIN HET-E-1.

GN HET-E-1.

OS Podosporea anserina.

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Lasiosphaeriaceae; Podosporea.

OX NCBI_TaxID=5145;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96009891; PubMed=7557402;

RA Saube S., Turco B., Begueret J.;

RT "A gene responsible for vegetative incompatibility in the fungus
 RT Podosporea anserina encodes a protein with a GTP-binding motif and G
 RT beta homologous domain.";

RL Gene 162:135-139(1995).

CC -1- FUNCTION: RESPONSIBLE FOR VEGETATIVE INCOMPATIBILITY THROUGH
 CC SPECIFIC INTERACTIONS WITH DIFFERENT ALLELES OF THE UNLINKED GENE.

CC HET-C.

CC -1- SIMILARITY: CONTAINS 10 WD REPEATS (TRP-ASP DOMAINS).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: L28125; AAA85775.1; -

DR InterPro: IPR001680; WD40.

DR Pfam: PF00400; WD40. 10.

DR PRINTS: PR00320; GPROTEINBPT.

DR SMART: SM00320; WD40. 10.

DR PROSITE: PS00678; WD_REPEATS.1; 10.

DR PROSITE: PS50082; WD_REPEATS.2; 10.

DR PROSITE: PS50294; WD_REPEATS_REGION. 1.

KW GTP-binding; Repeat: WD repeat.

FT NP_BIND 300 307 GTP (POTENTIAL).

FT REPEAT 839 869 WD 1.

FT REPEAT 881 911 WD 2.

FT REPEAT 923 953 WD 3.

FT REPEAT 965 995 WD 4.

FT REPEAT 1007 1037 WD 5.

FT REPEAT 1049 1079 WD 6.

FT REPEAT 1091 1121 WD 7.

FT REPEAT 1133 1163 WD 8.

FT REPEAT 1175 1205 WD 9.

FT REPEAT 1217 1247 WD 10.

SEQUENCE 1356 AA; 149765 MW; 965FB319844E0651 CRC64;

Query Match 15.0%; Score 455.5; DB 1; Length 1356;

Best Local Similarity 26.0%; Pred. No. 1.6e-26;

Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;

OY 28 EPPKRIIEKSLKQTNISCAR-LCLAOEYVCLASTAMK-----TENCVAKTKLANSTSS 81

DB 561 EMPDSDPDDPSLEIIVLCGSLFIRIRIVYFVHOSAKDFLLGTASDKSNKASQPAFE 640

OY 82 MIVPK-----QRLSASYEKEK-LCVKVFQDMSDVEVEVH--ISQNGHY 127

DB 641 LVPTGIEDSYIIFFMSLVNMSOKLRDITICLNAEGFLIDNVRVDPDPLATVRYSCIT 700
 OY 128 QHGINSYLKP-----MLQD-----FIT-----ALPARGIDH--AEN 159
 DB 701 WIDLRLDVSTSSKRWHLLODDGDIRFLTYLYLWLEALSLRALP-EGINAIQLES 759
 OY 160 ILSYLDKAKSLCAELVCKEYRYTSQGMIMKKILIER---WVRDSIMRGLAERKGGOTL 216
 DB 760 LFGHTIGRLIA--IVRQYR--ALSTRMIEKAPLOAVYSAL-----V 800
 OY 217 FKNKPPDGNAPNSFYRALY-----PKIIDIEETIESNMGRHSL----- 257
 DB 801 F-----APTSMIKIKKKEPGLISIVYEAEMNACQTLGHSVSVAFASD 852
 OY 258 -QRIHCS-----ETSG-----VYCLQY--DDQIVSGLRDNTIKIMDK 294
 DB 853 GQVAVSGSDDKTIKIMDTASGTQTLGHSVSVAFSPDEREAVASGSDTKIKIMDA 912
 OY 295 NTLCKRILGHTGVSVCIDY--DERVITGSSDSTVYRVMDVNTGEMLTLIHCEAVLA 352
 DB 913 ASGCTQTLGHSGRVSVAFSPDQGVAVSGSDHTIKIMDAAGCTQTLGHSVSVLS 972
 OY 353 LREN-NGMVTC-SKDRSLAVMDMASPTDITLRLVGRRAVNVDF--DKTIYVASG 408
 DB 973 VAFSPDQGVAVSGSDTKIKIMDTASG--TQTLGHSVSVAFSPDQGVAVSGSD 1029
 OY 409 DRTIKVNTSTCEFVRLNGHKKGIACIQYR--DLVYSSSSDNTLRMDIEGACLRVL 466
 DB 1030 DKTIKIMDTASGTQTLGHSVSVAFSPDQGVAVSGSDHTIKIMDAAGCTQTL 1089
 OY 467 EGHLEIVRCIRF--DNKRIYSGAVYDKIKVMDLVAALDPAPAGTLCITLVHSGRVR 524
 DB 1090 EGHGDSVSVAFSPDQGVAVSGSDITGIKIMD-----AASGT-CITLGHSGWVHS 1140
 OY 525 LQF--DEFQIVSSHDDTLIMDPLNDPAAQ 553
 DB 1141 VAFSPDQGVAVSGSDITGIKIMDAAGCTQV 1171

RESULT 13

POP1_SCHPO STANDARD; PRT: 775 AA.

AC P87060;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE WD-REPEAT PROTEIN POP1.

GN POP1 OR SPBC262.18.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RX MEDLINE=97347242; PubMed=9203581;

RA Kominami K., Toda T.;

RT "Fission yeast WD-repeat protein pop1 regulates genome ploidy through
 RT ubiquitin-proteasome-mediated degradation of the CDK inhibitor Rum1
 RT and the S-phase initiator Cdc18.";

RL Genes Dev. 11:1548-1560(1997).

SEQUENCE FROM N.A.

RP STRAIN=972;

RA Wood V., Rajandream M.A., Barrell B.G., Devlin K., Churche C.M.;

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN MAINTENANCE OF PLOIDY THROUGH PROTEASOME
 CC DEPENDENT DEGRADATION OF CDK INHIBITOR RUM1 AND S-PHASE INITIATOR
 CC CDC18. FUNCTIONS AS A RECOGNITION FACTOR FOR RUM1 AND CDC18, WHICH
 CC ARE SUBSEQUENTLY UBIQUITINATED AND TARGETED TO THE 26S PROTEASOME
 CC FOR DEGRADATION.

-1- SUBUNIT: BINDS TO CDC18.

```

CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y08391; CA69671.1;
CC EMBL: AL022103; CAAL789.1;
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00646; F-box; 1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PR00320; GPROTEINBRPT.
CC SMART: SM00256; FBOX; 1.
CC SMART: SM00320; WD40; 6.
CC PROSITE: PS50181; FBOX; 1.
CC PROSITE: PS00678; WD_REPEATS_1; 3.
CC PROSITE: PS50082; WD_REPEATS_2; 4.
CC PROSITE: PS50294; WD_REPEATS_REGION; 1.
CC Repeat; WD Repeat.
CC KW DOMAIN 298 345 F-BOX.
CC FT REPEAT 444 472 WD 1.
CC FT REPEAT 484 538 WD 2.
CC FT REPEAT 575 603 WD 3.
CC FT REPEAT 615 645 WD 4.
CC FT REPEAT 657 687 WD 5.
CC SEQUENCE 775 AA; 87816 MW; B06EDBA46553EEC1 CRC64;

Query Match 14.9%; Score 453; DB 1; Length 775;
Best Local Similarity 25.8%; Pred. No. 1.2e-26;
Matches 162; Conservative 86; Mismatches 229; Indels 150; Gaps 20;

QY 9 QEALFEMNSEREDC-----NNGEPRKIIPEKNSLRQTV-NSCARLCLNQET 56
DB 121 REKCLRRNSLSNLHANKRFLFNSGDSGNKKKEFTPTNTSNVVEYPPNCKSEKVASST 180
QY 57 VC-----LASTAMKTENCVAATKLANGTSSKATVKKRKLAA----- 92
DB 181 TFSLDAPNNSVANSYSPNLLGDSKTRQSPFPHSSSSSHNSLHEBYIVDFSENPISHP 240
QY 93 -----SYEKREL-----CYKFEQMSDQVEFEHLISQCHQGHINSYKLPMLOROF 144
DB 241 SNHSSQKNAVLLKLOLISFEKLPESVROYLLPHLSRCGRKAVONIRKILLPIFOKNP 300
QY 145 ITALPARGLDHIAENILSYLDALSCAELVCKEWTYR-VTSDGMLMKT-----I 193
DB 301 LTGPAA-----ETNLVYLTHDAPSLCAVSGVSHHWYKLVSSNELMKSLEFLKGFPMDSI 356
QY 194 ERVAVRDSLKRGIAERGMQOYLFKKRPPGNA-PNPSFYALYPKIIODETIESMWRG 253
DB 357 DSKIRFMCLRQSLA-----CAIKRYFERHFNLEERW--- 389
QY 254 RHLSTORHCHSESTKGYVC---LOYDDOKIVSGLRDNTIKIMDKNLECKRIILGHTGSV 310
DB 390 LHAPEKIKRCSFPIHGRLTKIOLFDDDKITIVSTCSPRINITYTKGVILRSLEEHGDV 449
QY 311 LCLQYDERVITGSSDSSTVAVWVDVNGEMLNTLIHCEAVLHLFPNGMVAWVCSKDRSIA 370
DB 450 WTEFYVDTLVSTGSTRVWMDLRTGCEKQVYFGHTSTIRCIKIVOGNSTTDTD----- 505
QY 371 VMDASPTDITLRVAVGHAANVVDFOCKYIVASGSDITIVWN----- 416
DB 506 -----DVEKE-----NRPASDANSMPYIISSRDCTIRLMSLPCLDPPPVANVE 552
QY 417 -----TSTCEVRLTNGHK---RGIAQLQYDRILVAVSGSSDNTIRLMDIEGAC 462
DB 553 NPQDNDFTSATITPFIYIRILRGHTDSYREVACL---GDLIVASVDGTLRVAKASTGYC 609

```

```

QY 463 LRVIEGHEELVRCIRPDNKR--IVSGAYDGKIKVMDLVALDPRAPAGTCLRTLYVHSG 520
DB 610 LHLVLRGHGRVYSTVITNPSRQOCISAGTADAKIRIMNL-----ESGEL-LQTLHGSHN 660
QY 521 RVFRLQFDEFOIVSSS--HDDTILIMD 545
DB 661 LVSOVTFNQNTIVASAPDPTSLRWMD 687

RESULT 14
CC4_YEAST STANDARD; PRT; 779 AA.
AC P07834;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CELL DIVISION CONTROL PROTEIN 4.
GN CDC4 OR YFL009W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88011240; PubMed=3309335;
RA Yochem J., Byers B.;
RT "Structural comparison of the yeast cell division cycle gene CDC4 and
RL J. Mol. Biol. 195;233-245(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RT "Analysis of the nucleotide sequence of chromosome VI from
RL Saccharomyces cerevisiae."
RN [3]
RP Net. Genet. 10:261-268(1995).
RN [3]
RP SEQUENCE OF 1-579 FROM N.A.
RC STRAIN=S288C / AB972;
RA Barrell B.G., Churcher C., Rajandream M.A.;
RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN IS ESSENTIAL FOR INITIATION OF DNA
CC REPLICATION AND SEPARATION OF THE SPINDLE POLE BODIES TO FORM THE
CC POLES OF THE MITOTIC SPINDLE. IT ALSO PLAYS A ROLE IN BUD
CC DEVELOPMENT, FUSION OF ZYGOTIC NUCLEI AFTER CONJUGATION AND
CC VARIOUS ASPECTS OF SPOKULATION. REQUIRED FOR HTAL-HTBI LOCUS
CC TRANSCRIPTION ACTIVATION.
CC -1- SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
CC -1- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X05625; CA29113.1;
CC EMBL: D50617; BAA09229.1;
CC EMBL: D31600; BAA06495.1;
CC EMBL: 246255; CA86341.1;
CC PIR: A26867; A26867.
CC SGD: S0001885; CDC4.
CC InterPro: IPR001810; F-box.
CC InterPro: IPR001680; WD40.
CC Pfam: PF00646; F-box; 1.
CC Pfam: PF00400; WD40; 7.
CC PRINTS: PR00320; GPROTEINBRPT.

```


DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 4.
 DR PROSITE; PS00682; WD_REPEATS_2; 5.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Cell division; Mitosis; Sporulation; Repeat; WD repeat.
 FT DOMAIN 272 319 F-BOX.
 FT REPEAT 380 408 WD 1.
 FT REPEAT 420 449 WD 2.
 FT REPEAT 461 493 WD 3.
 FT REPEAT 528 556 WD 4.
 FT REPEAT 568 598 WD 5.
 FT REPEAT 630 658 WD 6.
 FT REPEAT 669 698 WD 7.
 FT CONFLICT 460 460 K -> E (IN REF. 1).
 SQ SEQUENCE 779 AA; 86089 MW; 0348F2F8FA76F3BC CAC64;

Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 1.4e-22;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

OY 56 TWCLASTAKTENCYAKFTLANGSTSMIVPKRKLSASYEK-----ELCYKYEQW 108
 DB 194 TTPPLAKTYTINN-----NNNIADLESKDSITISPEYLSDEIFSAINNMPHAYFK- 244
 OY 109 SSSDOVEFEHLISOMCHYOHGHSYKPMLODFITALPARGDHI AENITSLYDAKS 168
 DB 245 -----NLLFRIVAMNDSSELSDELTLIKDNKRLITSLP-----ELSLKFNLYLOED 294
 OY 169 LCAAEIVCEKWR-VTSDGMWKLI--ERWYRDSLWRGLAERRGWQYLFFKNKPPDGN 225
 DB 295 IINSVSQNMWKILRKSTSLMKLLISENFV-----SPKGF 331
 OY 226 APPNSFYRLVYKIID-----LET--ISNMRCGRHSIORICRSETSKGYCILOYD 276
 DB 333 NSLNLKLSQKPKLSQOORLISFLENFLLKNMYPFVQRTTLRGHMSVITCLOFE 391
 OY 277 DQIVSGLRDNTIKIMDKTECKRILGHTGVSYLQLOYDE-RVLIITGSSDSTVRWVDN 335
 DB 392 DNYVITGADDKMIRYDSINKKFLQLSGHDGVAWLKVAAGILVSSSTRTYRWMDIK 451
 OY 336 TGEMLNTLIHCEAVLHLRFNNGMAYVTSKRSIAVWDMASPTDITLRLVVGRAAVNV 395
 DB 452 KG-----CCT-----HYFKGHNSTVRC 468
 OY 396 VDFDD-----KIYVASGDRITKVMNT-----STCE-----FVRLING 428
 DB 469 LDIVYKNIKIKIYVGSKRNTHLWKLPESSVPPDHGEHDYPLVPHPEENPYFVGVLRG 528
 OY 429 HKRGIAQLQYDRVLVSGSSDNTIRLMDIEGACLRVLEGHLELVRCIREDN--KRIYSG 486
 DB 529 HMAVRYTSGHGNIVSGSYDNTLIVMDVQMKCIYLSGTHDILYSIYHKKRCSA 588
 OY 487 AYDGKIKYWDL-----VAALDPRAPATL--CLRTLVHSGRVPRLQDFEQVSS 535
 DB 569 SMDTTIRIWDLENIMWNGECYATNSAPCAKILGAMVYLOGHTALVGLLRLSDKFLVSA 648
 OY 536 SHDDTILIMDFLPPAAQAEPSPRSRTYTY 566
 DB 649 AADGSIRGWD-AND-----YSKRFST 668

RESULT 15
 POP2_SCHPO STANDARD; PRT; 703 AA.
 AC 014170;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE WD-REPEAT PROTEIN POP2 (PROTEOLYSIS FACTOR SUD1).
 GN POP2 OR SUD1 OR SPAC4D7.03.

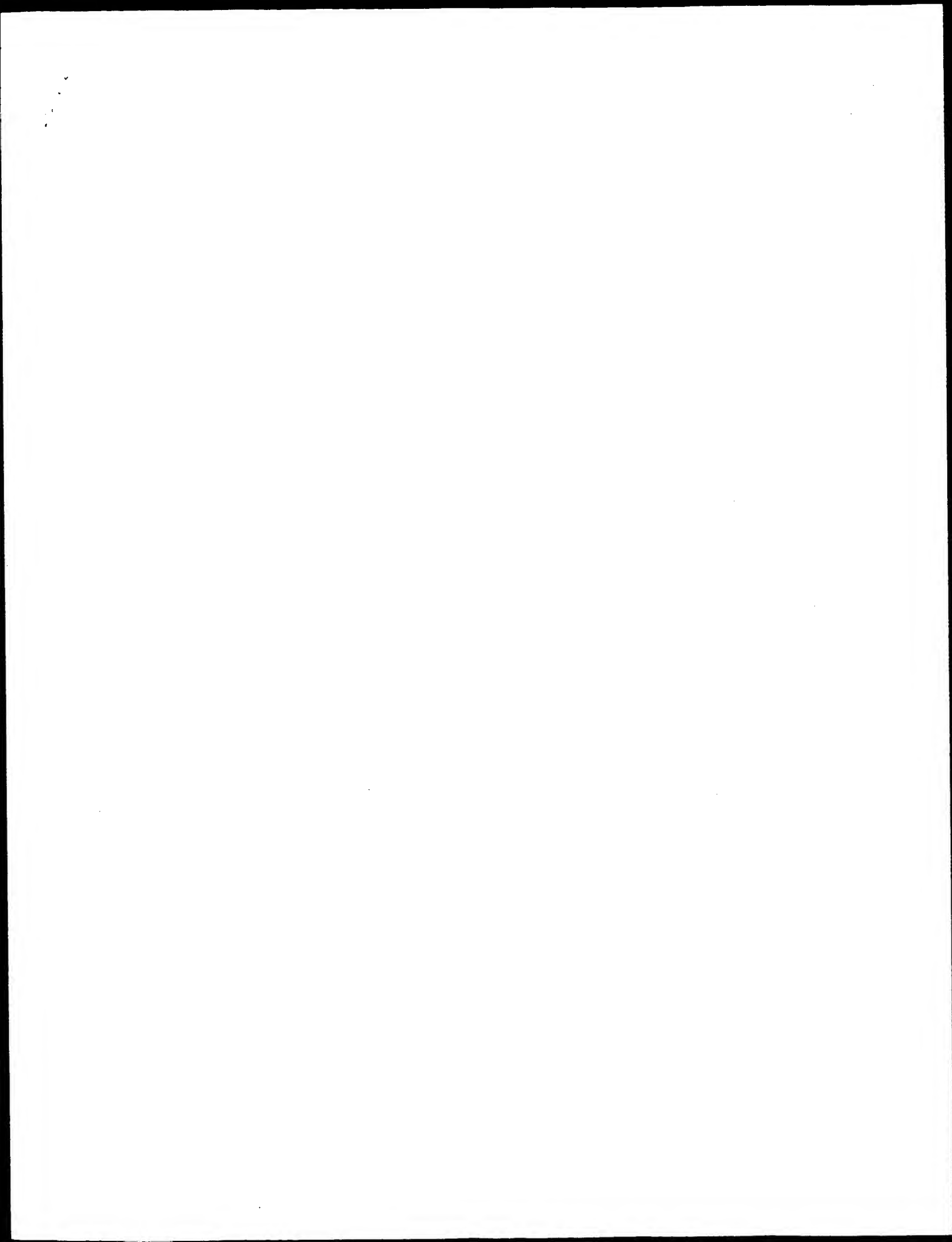
OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN 1;
 RC SEQUENCE FROM N.A.
 RP STRAIN=972;
 RL wolf D.A., Jackson P.K.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN 12;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98318628; PubMed=9653157;
 RA Jallepalli P.V., Tien D., Kelly T.J.;
 RT "sud1 targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1
 RT proteins for degradation and stops unwanted diploidization in fission
 RT yeast.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8159-8164(1998).
 RN 13;
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Gentiles S., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: PUTATIVE CYCLIN-DEPENDENT OF E3 (PROTEIN-UBIQUITIN LIGASE)
 CC COMPLEX: TARGETS CYCLIN-DEPENDENT KINASE-PHOSPHORYLATED CDC18 AND
 CC RUM1 PROTEINS FOR DEGRADATION.
 CC - SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
 CC - SIMILARITY: CONTAINS 1 F-BOX DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AF03867; AAC95480.1; -
 DR EMBL; AF064515; AAC39496.1; -
 DR EMBL; Z98602; CAB11275.1; -
 DR InterPro; IPR001810; F-box.
 DR InterPro; IPR001680; WD40.
 DR Pfam; PF00646; F-box; 1.
 DR Pfam; PF00400; WD40; 7.
 DR PRINTS; PR00320; GPTEINBRPT.
 DR SMART; SM00256; FBOX; 1.
 DR SMART; SM00320; WD40; 6.
 DR PROSITE; PS00181; FBOX; 1.
 DR PROSITE; PS00678; WD_REPEATS_1; 3.
 DR PROSITE; PS00682; WD_REPEATS_2; 6.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 FT DOMAIN 236 283 F-BOX.
 FT REPEAT 369 417 WD 1.
 FT REPEAT 429 473 WD 2.
 FT REPEAT 505 533 WD 3.
 FT REPEAT 545 575 WD 4.
 FT REPEAT 587 615 WD 5.
 FT REPEAT 625 654 WD 6.
 SQ SEQUENCE 703 AA; 79605 MW; 0336A0566C152396 CAC64;

Query Match 13.1%; Score 396.5; DB 1; Length 703;
 Best Local Similarity 23.3%; Pred. No. 1.9e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

OY 14 KFNNSERED--CNGGEPRKT-----IPEKSLKQTYNSCARLCL---NOETVCL 59
 DB 96 KFNESLKNYSYKCLDHPDSVPGDFSSILVQPRNFL-YSHSLPCKIISIDRRNRKILD 154
 OY 60 ASTAMKTECYAKTKLANGSTSMIVPKRKLSASYEK-EKELCYKYEQWMSQVEVE 118
 DB 155 NTSISNSDNFPSPKPV--DTSNTVSPGSKPISEDLIDNLSIVOTFEDLDEGIO-STAF 211

QY 119 HLIQOMCHYQHG-INSTLAPMLQDFTTALPARGLDHIAENILSTYDANKSLCAELVCK 177
Db 212 FOLLRSQNRQSMRLLECEPELKKDILSNLP-----SIVQSLINDHISPLSCRIVSP 267
QY 178 EMYRV-----TSDGMLTKKLEJRMVWRTSLWRGLAERGWQOYLEFKNPPDGNAP-----N 229
Db 268 TWRNILDVHTS---YMKHMF-----SLFGFOINENDW-----KYANPNLNRPPPLHND 312
QY 230 SFYRALYPKIIDIETIESNMWRCGRHSIOR----- 259
Db 313 QISDDYFPELIFK-----RHFLENKRWLFPSPISPHLSFPIHVENEMITSLHL 360
QY 260 -----IHCSEYISK-GVYCLQYDDQKIVSGLRDNTIKIMDKNT 296
Db 361 KDRITTTSSGCTIQIHNAITGVLEARLEHKEGVMAVKIHENTLVSGSIDKIVRWNNIEK 420
QY 297 LECKRIITGHTGSLVLCIY-----DERVITITSSDSTVWMDVNTGEML 340
Db 421 AKCTHIFRGHISIRCELEIVPSRLIRHGEIYEPDQPIYVSGSRDHTLRW----- 472
QY 341 NTLIHCEAVLHLRFNNNGMMWTCCKDRSIAVWDMASPTDITLRVLVGHRAAVNVVDFDD 400
Db 473 -----KLPKNTDPPYLPDNTNSIDRME-KNPYFV---HTLIGHDSVRTISGYG 517
QY 401 KYIVSASGDRITIKVWNTSTCEFYRTLNGHKGRTACIQYRD--RLVYSGSSDNTIRLMDIE 458
Db 518 DIIYSGSYDSSIKIMRVSTGECCLYHLRGHSLRIYSVLYEPERNICISGSMDSIRWMDLS 577
QY 459 CGACLVLEEGHEELVRCIRFDRNRIVSGAYDGIKIKWDLVAALDPAPAGTLCIFLVEH 518
Db 578 TGTCKYVLEEGHDAFVTLVNFQNRILISGSADSTIRIMD---LNTGKP-----LMVLPSPN 628
QY 519 SGRVFRLOFDEPOIVSSHDDTILMD 545
Db 629 SGYISSFVSDEHKIT-SGNDGSVKIWD 654

Search completed: May 8, 2002, 10:54:21
Job time: 213 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 8, 2002, 10:49:08 ; Search time 19.54 Seconds
(without alignments)
2218.184 Million cell updates/sec

Title: US-09-601-168a-2

Perfect score: 3034

Sequence: 1 MDPAAVLQERKALKFNSSSE.....PAAQAEPPSPSRITYISR 569

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2597	85.6	518	2 B48088	beta-transducin re
2	1635.5	53.9	701	2 T16607	hypothetical prote
3	690	22.7	506	2 T50211	WD-repeat protein
4	590.5	19.5	605	2 T38832	probable sulfur me
5	545	18.0	650	2 S49932	MET30 protein - ye
6	531.5	17.5	650	2 T46660	sulfur controller-
7	520	17.1	579	2 T22703	hypothetical prote
8	519.5	17.1	267	2 S62507	hypothetical trp-a
9	455.5	15.0	1356	2 T18521	beta transducin-11
10	453	14.9	775	2 T45136	WD repeat protein
11	399	13.2	779	2 S56245	cell division cont
12	396.5	13.1	703	2 T43557	F-box/WD-repeat pr
13	354	11.7	409	2 S36113	Lis-1 protein - hu
14	354	11.7	410	2 S48052	platelet-activatin
15	341	11.2	515	2 S19487	hypothetical prote
16	339.5	11.2	777	2 T41075	hypothetical WD-re
17	337	11.1	1466	2 A55532	myosin-heavy-chain
18	332.5	11.0	317	2 T46032	WD-40 repeat regul
19	330	10.9	333	2 G85034	probable WD-repeat
20	318	10.5	422	2 A56640	CDC4 repeat unit-c
21	313.5	10.3	376	2 T19265	hypothetical prote
22	312.5	10.3	714	2 S56893	hypothetical prote
23	312	10.3	876	2 T51507	WD40-repeat protei
24	307.5	10.1	1693	2 S76086	beta transducin-11
25	305	10.1	704	2 S33263	transcription init
26	299.5	9.9	1194	2 T03818	apoptotic proteina
27	297	9.8	659	2 T41148	trp-asp repeat con
28	292	9.6	659	2 S38108	hypothetical prote
29	290.5	9.6	614	2 S58306	WD-40 repeat regul

30	287.5	9.5	327	2 S48839	guanine nucleotide
31	285.5	9.4	798	2 S34023	TATA box-binding p
32	283	9.3	713	2 JN0133	WD-40 repeat regul
33	280	9.2	494	2 T19550	hypothetical prote
34	278.5	9.2	380	2 T40283	beta-transducin -
35	277	9.1	586	2 T38992	WD-40 repeat regul
36	276.5	9.1	501	2 T27512	hypothetical prote
37	274	9.0	2241	2 T16064	hypothetical prote
38	273	9.0	325	2 T09613	probable GTP-bind
39	272.5	9.0	1008	2 T32986	hypothetical prote
40	272.5	9.0	1019	3 JC7538	neuronal different
41	271.5	8.9	817	2 S51445	probable membrane
42	270	8.9	451	2 S65162	hypothetical prote
43	270	8.9	473	2 T33805	hypothetical prote
44	270	8.9	504	2 T50983	probable pleiotrop
45	270	8.9	651	2 T50289	WD repeat protein

ALIGNMENTS

RESULT 1

B48088 beta-transducin repeat-containing protein - African clawed frog

N.Alternate names: beta-trcp

C.Species: Xenopus laevis (African clawed frog)

C.Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000

C.Accession: B48088

R.Species: W.; Keiper, B.D.; Stralowa, C.; Castanon, M.J.

Mol. Cell. Biol. 13, 4953-4966, 1993

A.Title: Saccharomyces cerevisiae cdc15 mutants arrested at a late stage in anaphase

A.Reference number: B48088; MUID:93330289

A.Accession: B48088

A.Status: preliminary

A.Molecule type: mRNA

A.Residues: 1-518 <SP>

A.Cross-references: GB:M88268; NID:9295542; PIDN:AAA02810.1; PID:929543

C.Superfamily: unassigned WD repeat proteins; WD repeat homology

C.Keywords: duplication

F:431-462/Domain: WD repeat homology <WD1>

Query Match	85.6%	Score 2597	DB 2	Length 518
Best Local Similarity	91.6%	Pred. No. 1.6e-193		
Matches 488	Conservative 7	Mismatches 8	Indels 30	Gaps 1
QY	18	SSEREDCNNGEPPRKITPEKNSLRQYNSCARLQNETVCLASTAMKTEVCVAKTKIAN	77	
DB	13	ASEREDCNNDPEPKITTEKNTLRQ-----TKIAN	42	
QY	78	GTSSMIVPKORKLSASYEKELCVKYEQWSESDQVEFVEHLISOMCHYOHGINTYK	137	
DB	43	GTSSMIVPKORKLSANYEKELCVKYEQWSESDQVEFVEHLISOMCHYOHGINTYK	102	
QY	138	PMQGRFTALPARGLDHIENILSYLDAKSLCAELVCKEYRYSQGLMKKLEIRV	197	
DB	103	PMQGRFTALPARGLDHIENILSYLDAKSLCAELVCKEYRYSQGLMKKLEIRV	162	
QY	198	RIDSILMRLAERGMQYLFKNKPPDGNAPNSFFRALYPKTIQDIETESNMRCGRSL	257	
DB	163	RIDSILMRLAERGMQYLFKNKPPDGNAPNSFFRALYPKTIQDIETESNMRCGRSL	222	
QY	258	ORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKINDKNTLECKRILTGHTGVSCLQYDE	317	
DB	223	ORHCRSETSKGYVCLQYDDOKIVSGLRDNTIKINDKNTLECKRILTGHTGVSCLQYDE	282	
QY	318	RYITIGSSDSTYRVMDVNGEMLNLTLLHHCFAVLHFRNGMAYCSKDRSLAVMDMSP	377	
DB	283	RYITIGSSDSTYRVMDVNGEMLNLTLLHHCFAVLHFRNGMAYCSKDRSLAVMDMSP	342	
QY	378	TDITLRLVLYGHRANVNVDFDDKYIVSASGDRITIKVWNTSTCEFRRLNGHKRGIAQLQ	437	
DB	343	TDITLRLVLYGHRANVNVDFDDKYIVSASGDRITIKVWNTSTCEFRRLNGHKRGIAQLQ	402	

QY 438 YRDLRVSSSSDNTIRLMDIEGACIRYLEGHEELVRCIRPNKRITVSGAYDGKIKVMDL 497
 Db 403 YRDLRVSSSSDNTIRLMDIEGACIRYLEGHEELVRCIRPNKRITVSGAYDGKIKVMDL 462
 QY 498 VAALDPRAPAGTLCITLVEHSGRVFRLQDFEFOIVSSSHDITLIMPLNDP 550
 Db 463 VAALDPRAPAGTLCITLVEHSGRVFRLQDFEFOIVSSSHDITLIMPLNDP 515

RESULT 2
 116607
 Hypothetical protein K10B2.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16607
 R:Miller, N.
 submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid K10B2.
 A:Reference number: 218545
 A:Accession: T16607
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-701 <MIT>
 A:Cross-references: EMBL:U28730; NID:9860694; PID:9860695; PIDN:AAA68258.1; CESP:K10B2.1
 A:Experimental source: strain Bristol N2
 A:Gene: CESP:K10B2.1
 A:Introns: 78/3; 125/1; 183/2; 281/3; 404/3; 551/3; 668/3

Query Match 53.9%; Score 1635.5; DB 2; Length 701;
 Best Local Similarity 57.7%; Pred. No. 8.6e-119;
 Matches 328; Conservative 69; Mismatches 116; Indels 55; Gaps 8;

QY 31 RKIIKESKSLRQTYNSCARLQINQETVCLASTAMKTENCVAKTKIAN-----GTS 80
 Db 2 RFRFGKRLALQGRAROGSGTQALTYCVST---TERCF--TAVSNPIFLFTSTFVSF 55
 QY 81 SKIYVKQR-----KLISAYEKEKELCVKYFEQWSESQDVEFEHLISQMHYQGHINSY 135
 Db 56 SFLPFSRNTQIFLTSYRSFSFSEVL-----KMSHEQDLDMKIVIRHLSHYQGVKNDF 110
 QY 136 LKPMIQRPTALPARGDHTAENLSTYLDAKSLCAELVCKEYRRTSDGMKKLIER 195
 Db 111 IRPMIQRPTISNLP-----HLVELLLENVSDLSKSEVSTSRCALARGQHMKKLIER 166
 QY 196 MYRTDSLWRGLAERGMGYL-----FKKKPPDGNAAPNSFYRAL 235
 Db 167 MYRSDSLWGLSEKRWOMKFLNISRDMSVYRICEKFNVDVAKIKDKLDQLIMHVFYSKL 226
 QY 236 YPKTIQDIETESMWRGCRSLORIHCRSETSKGVCIQYDDOKITVSGLRDNTIKIMDKN 295
 Db 227 YPKTIIRDHNIDNMWRKRNMYKTRINQOSENSKGVCIQYDDOKITVSGLRDNTIKIMDRK 286
 QY 296 TLECKRIHTGTVLCQYDERVLIITGSSDSTVRVMDVNTGEMNLNLHHCFAVLHRT 355
 Db 287 DYCGRILISGHTSVLCQYDNRVLIISGSDATVRVMDVEGDEIKTLHHCFAVLHRT 346
 QY 356 NGGMVNTCSKDRSLAVNDMASPTDITLRLVGLHRAAVNVVDEPDOKITVSGASGRTIKV 415
 Db 347 ANGIMVNTCSKDRSLAVNDMASPTDITLRLVGLHRAAVNVVDEPDOKITVSGASGRTIKV 406
 QY 416 NTSTCEVFTLNGHKGRIAGLCQYRDLRVSSSSDNTIRLMDIEGACIRYLEGHEELVRC 475
 Db 407 SMDTLFEVFTLAGHRIGIACQYRDLRVSSSSDNTIRLMDIHSQVCLRVLEGHEELVRC 466
 QY 476 IRPNKRITVSGAYDGKIKVMDIVALDPRAPAGTLCITLVEHSGRVFRLQDFEFOIVSS 535
 Db 467 IRPEKRTIVSGAYDGKIKVMDIVALDPRAPAGTLCITLVEHSGRVFRLQDFEFOIVSS 526
 QY 536 SHDITLIMPLNDPAAQAPRPSPT 563

Db 527 SHDITLIMPLNDP-----PSGLPST 549

RESULT 3
 T50211
 WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 02-Sep-2000
 C:Accession: T50211
 R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Brown, S.; Murphy, L.; Jones, L.;
 submitted to the EMBL Data Library, January 2000
 A:Reference number: 225046
 A:Accession: T50211
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-506 <MCD>
 A:Cross-references: EMBL:AL136538; PIDN:CAB66464.1; GSPDB:GN00066; SPDB:SPAC30.05
 A:Experimental source: strain 972h(-); cosmid c30
 C:Gene: SPAC296.01; SPDB:SPAC30.05
 A:Map position: 1
 A:Introns: 43/1; 74/3
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 22.7%; Score 690; DB 2; Length 506;
 Best Local Similarity 30.4%; Pred. No. 1.1e-45;
 Matches 163; Conservative 104; Mismatches 167; Indels 102; Gaps 15;

QY 67 ENCVAR---TKLANGTSMIYVKQRKLSAYEKEKELCVKYFEQWSESQDVEFEHLIS 122
 Db 8 KNVSVSVSLTSCSDSTSPVCLNPLS-----HENNRIDLRLDLA 50
 QY 123 QMCHYQGHINSYKRLMDRDTLALPARGLDHTAENLSTYLDAKSLCAELVCKEYR 182
 Db 51 SLSEGVAVYNNHVRSLTDFEVPF---EVSLSRVSYLDQLDLCCKIMSRMRL 106
 QY 183 TSDGMLMKLI-----ERNVRTDSLWRG-----LAERGMG----- 213
 Db 107 LEDPGIMKALYQMGFVNEVNLNEFEAMRRTNHPQPFENFLQONITIGYGMFLPQ 166
 QY 214 QYLFKKPPDGNAAPNSFYRALYPKTIQDIETESMWRGCRSLORIHCRS----- 264
 Db 167 QYTF-----DSMGRPLNMSYLY---KEHAHLSWRRGRLVSTFNPSIRPADDF 217
 QY 265 -ETSKGVLCQYDDOKITVSGLRDNTIKIMDKNTLCEKRLITGHTSVLCQYDER--VII 321
 Db 218 RATLDSVYCVQYDDDELIMVSGSDRTYVMDVNSRFLIKLYHSGSVLCIDPCCRRLIV 277
 QY 322 TGSSDSTVRVMDVNTGEMNLNLHHCFAVLHRTFNNGMVTCSKDRSLAV--DNASPTD 379
 Db 278 SGSSDSTIITIMQNRPLKVFEGHTDNLGVVSENYIISSRDHTARVRLDATSPA 337
 QY 380 IILRLVGLHRAAVNVVDEDDK--YIVSASGRTIKVNTSTCEVFTLNGHKGRIAGLCQ 437
 Db 338 ACM-HYLRGHLASVNSVQSKTGLITVASSDRTLRMTDITTYGHICIRIIAHORICAO 396
 QY 438 YRDLRVSSSSDNTIRLMDIEGACIRYLEGHEELVRCIRPNKRITVSGAYDGKIKVMDL 497
 Db 397 YNGKFTVSSSSDNTIRIPFASGKILRLHOGHEDLIRYVRNDEYIVSGGIDGYRILN- 455
 QY 498 VAALDPRAPAGTLCITLVEH-----SGRVFRLQDFEFOIVSSSHDITLIMDF 546
 Db 456 -----FNTGEQHCVLNHSRNSRVFGIQDFHRRILACTHSEILVNNF 497

RESULT 4
 T38932
 probable sulfur metabolite control protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: T38932
 R:Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.

A:Cross-references: EMBL:U17251; NID:g806757; PIDN:AAA68968.1; PID:g806758
 C:Genetics:
 A:Gene: scon-2
 A:Map position: 3
 A:Introns: 75/3; 319/1; 354/1
 C:Function:
 A:Description: negatively regulates sulfur structural gene expression
 A>Note: scon-2+ expression is dependent on CYS3 function and the binding of CYS3 to the
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 17.5%; Score 531.5; DB 2; Length 650;
 Best Local Similarity 25.0%; Pred. No. 3,1e-33;
 Matches 144; Conservative 76; Mismatches 180; Indels 177; Gaps 11;

QY 117 VEHLSQKCHYOHGHNSTLKPMLORDFTTALPARGDLHIAENILSYLDASLCAELVC 176
 Db 99 LOGILSOLCOFFPOLSFVSREVNEMALKIDFISALPV---ELAQKVCYLDYVSLKRAQVS 154
 QY 177 KEMVRYTSDGMKMLKILRMVRTDLSLGRGLAERGMQGYLFKKNP-----P 222
 Db 155 QMWRILASDAVAVWVRMCDHVNRR-----CTKCGMGLPLERKKLRNRYRQOLAKGP 208
 QY 223 DQNA-----PNSFYRA 234
 Db 209 QGRVTELDASHDSQDRSVNQHGRPAEAEEEDPIKRCQMAAAEASKAVYQPTRSKKA 268
 QY 225 LYPKTIQDIETESNWRGGRHSIORHCRSETSKGVCLQYDQKIVSGLRDNTIKIWDK 294
 Db 269 YV---RDRWQVSYMKNSRYKLSVL---KGHEMGVTCLOQDDMLAIGSYDTIKIWN 321
 QY 295 NTECKRILITGHSVLCLOYDERIITNGSSDSTVYRWVDVTGEMNLNLHHCANVHLR 354
 Db 322 ETEECIRLVTGHTAGIRALQFDSDKLSLDHRIKWNHITGECISLFAAHTPSVSVH 381
 QY 335 FNGGMAVTCSDKRSIAVWDMASPTDITLRVLYGHRAAVNV---VDFDKYIVSASGRTI 412
 Db 382 FDGHLLASGSSDKTYKIDFNFSKEYCYLK---GHSQWVNSTHVDIKSRVYFSASDDTTI 437
 QY 413 KVMNTSTCEYRTLNGHKRGT-----ACLOYRDLRVS----- 445
 Db 438 KLMDLDFQVIRTEYEGHVQOVLILPPEYEPDEVLNGASQNDQAMSVSSGSGSPS 497
 QY 446 -----GSSD 449
 Db 498 MSHAQIERAGSPGSHSSSHNLLPSSLPDGEDVRLHYGSAFVADSEPRLPPEYMTGGLD 557
 QY 450 NTIRLMDIEGACLRVLEGEHELVCIRFDNKRIVSGAYDQKIKVMDLVAALDPRAPAGT 509
 Db 558 STMRLMDSATGRCLRTLFGLHEGYWSLAGDITRIVISGANDQMKYTW-----PRSG 608
 QY 510 LCLRLIVEHSGRVERLQDFEIVSSSHDDTILIMDF 546
 Db 609 KCDATYTGCHGCVTCVGLSDLSLMASGSEDGIRLHSF 645

RESULT 7
 T22703
 hypothetical protein F55B12.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T22703
 R:Simms, M.
 submitted to the EMBL Data Library, September 1996
 A:Reference number: Z19602
 A:Accession: T22703
 A:Status: preliminary; translated from GB/EMBL/DDSD
 A:Molecule type: DNA
 A:Residues: 1-579 <MIL>
 A:Cross-references: EMBL:Z79757; PIDN:CAB02129.1; GSPDB:GND0023; CESP:F55B12.3
 A:Experimental source: clone F55B12
 C:Genetics:
 A:Gene: CESP:F55B12.3

A:Map position: 5
 A:Introns: 22/3; 45/2; 77/3; 122/3; 171/2; 211/2; 342/2; 417/3; 478/3; 526/1

Query Match 17.1%; Score 520; DB 2; Length 579;
 Best Local Similarity 28.8%; Pred. No. 2e-32;
 Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

QY 72 KTKLANGTSSNI-----VPKORR--LSASYEKEKEKEL-----CVKTFEOMSESDOVE 115
 Db 27 ESSYSNCGSSSSYNADKLSSSRPLQHKLDLSASSRRNNDLPRVEHLIALFKDLSAEOQMD 86
 QY 116 FVEHLISQKCHYOHGHNSTLKPMLORDFTTALPARGDLHIAENILSYLDASLCAELVC 175
 Db 87 AFTRLQESNMNTNIROLRAIIEPHFORDFLSCLPV---ELQKILHNLGTGDLKVAQY 142
 QY 176 CKEMVRYTSDGMKMLKILRMV-----RIDSILRGLAERGMQGYLFKKNPDPGNAP 227
 Db 143 SKMVKLISEIDKTIKSLGYVEEFKHPDPDRTYTGAMOGTAIAAG-----VTIPDHQIP 195
 QY 228 PN-SFYRALYPKTIQDI-----ETIESNWRGGRHSIORHCRSETSKGVCL 273
 Db 196 CDLNVHRFLKLOKFGDIFERAADKSRYLRADKIEKMNANPIMGSAV-LRGHEDHVTICM 254
 QY 274 QYDQKIVSGLRDNTIKINDKNTLECKRILITGHSVLCLOYDE--RVITGSSDSYVRY 331
 Db 255 QIHDDVLYGSDNDNTLKWCIKGEVMTYLVGTGCVMTSQISQGRYIVSGSTDRYVY 314
 QY 332 WDVNTGEMNLNLHHCANVHLRFNNGMNVTCSDKRSIAVWDMASPTDITLRVLYGHR 391
 Db 315 WSTVDSGLHTLOGHTSYRCMAAGSILVTSGRDITLWVWVESGRHLA---TLGHHA 371
 QY 392 AVNVVDFDKYIVSASGRTIKVWNTSTCEFYRTLNGHKRIACILOYDR--LVVSSSD 449
 Db 372 AVRCVQDFDGTYYVSGGYDTYKIKMNAHTGRCIRTLGHNRRYSLLFEBERSIVSGSLD 431
 QY 450 NTIRLMDI---ECGACLRVLEGEHELVCIRFDNKRIVSGAYDQKIKVMDLVAALDPRAP 506
 Db 432 TSIRVWDFTRPEQECVALLQGHSTLSLQMGRLNIVSCNADSHVRWDI-----H 483
 QY 507 AGTLCRLIVEHSGRVERLQDFEIVSSSHDDTILIMDF 545
 Db 484 EGT-CVHMLSGHRSATITSLOWFGRNWVATSSDGTIVKMD 522

RESULT 8
 S62507
 hypothetical trp-asp repeat-containing protein - fission yeast (Schizosaccharomyces p
 C:Species: Schizosaccharomyces pombe
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 26-May-2000
 C:Accession: T38502; S62507
 R:Jones, L.; Murphy, L.; McNeil, A.; Simpson, I.; Harris, D.; Barrell, B.G.; Rajandre
 submitted to the EMBL Data Library, October 1995
 A:Reference number: Z21798
 A:Accession: T38502
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-267 <J02>
 A:Cross-references: EMBL:Z66525; NID:g1044926; PIDN:CAA91423.1; PID:g1044927; GSPDB:G
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:18-51/Domain: WD repeat homology <MIL>
 F:104-137/Domain: WD repeat homology <MD2>

Query Match 17.1%; Score 519.5; DB 2; Length 267;
 Best Local Similarity 38.0%; Pred. No. 7.5e-33;
 Matches 104; Conservative 58; Mismatches 83; Indels 29; Gaps 6;
 QY 266 DWTIKIMKNTLECKRILITGHTGVSVCLOYDER--VIITGSSDSYVRYWVDVTGEMNLTL 343
 Db 1 DRTVSWDVNRSFLLKLYGHSGSVLCIDFCRRRLNLYSGSSDSSTIIIMQNRRLPKVY 60

QY	344	IHCEAVLHLEAFNNGMATWCSDKSIAMV--DMSPPDITLRLRVLGHRAAVNVYDDDK	401
DB	61	FGHDNVLGVVVASENYIISSSRDHTARWRLDASPACM-HVLRGLHLSVNSVQYSSK	119
OY	402	--YIVASGDRPTIKVMWNTSTCEPFTLNLGHKRGJALQYRPRLVSGSSDDTILAMDIEC	459
DB	120	TGLIVTASDPTLRLWDTTGTGHCIRITHAHORGJACADQYNKPFIVSGSSDITIRIFEAAS	179
OY	460	GACLRVLEGHLEELVNCIRFDNKKRIIVSGAYDGKIKVMDLVAAIDPRAPAGTLCRTLVH-	518
DB	180	GKLRLMLQGHEDLITVAFRNDKIVSGGVDGTVRIWN-----FNTGEQHC	224
OY	519	-----SGRFRLODFEQIVSSSHDITILWDF	546
DB	225	VLHNSRNSRVFGLQDFDRRIIACHTSSSEILWNP	258
RESULT	9		
	T18521		
		beta transducin-like protein - Podospora anserina	
	C:Species:	Podospora anserina	
	C:Date:	15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999	
	C:Accession:	T18521	
	R:Saupé, S.; Turcq, B.; Begueret, J.		
	Gene 162, 135-139, 1995		
	A:Title:	A gene responsible for vegetative incompatibility in the fungus Podospora anserina	
	A:Reference number:	Z18944; MUID:96009891	
	A:Status:	preliminary; translated from GB/EMBL/DBJ	
	A:Molecule type:	DNA	
	A:Residues:	1-1356 <SAU>	
	A:Cross-references:	EMBL:L28125; NID:9607002; PID:9607003; PIDN:AA65775.1	
	A:Gene:	het-e1	
	A:introns:	761/3	
Query Match	15.0%; Score 455.5; DB 2; Length 1356;		
Best Local Similarity	26.0%; Pred. No. 6.9e-27;		
Matches 164; Conservative 102; Mismatches 220; Indels 145; Gaps 30;			
OY	28	EBPRKIIPKNSLRQTVNSCAR-LCLNQETVCLASTAMK-----TENCVAKTKIANGTSS	81
DB	581	EMPDVDSDPESLEIEIVKLGSGFLIIRERTVYFVHQSAKDFLGTASDKASNKASQEAFAE	640
OY	82	MYPK-----QRKLSASYEKEK--LCVKKFEQMSDSQVPEVHL--ISOMCHY	127
DB	641	IVFPTGIEDVSYIIIFKRSILNVMSQKLRDIYCLNAPGFLIDNVRPDPDPLATVRYSCIY	700
OY	128	HGHINISLYKP-----MLORD-----FIT-----ALPARGLDHI--AEN	159
DB	701	WIDHLDVIVSYSSSKWVHLLODDGDIDHFLITKYLWLEALSRLRLP-EGINALRQLES	759
OY	160	ILSYLDAKSLCAELVCEKMYRVTSQDMLMKLIER--MVRTDISLRGLAERGGQYL	216
DB	760	LIGHTIRGLIA---IVRDGYRP--ALSYRMIIIEKAPLQAVTSAL-----V	800
OY	217	FKMKPPDGNAPNSFPRALY-----PKIIDIEFTIESNMCGHSL-----	257
DB	801	F-----APDMSMIKKIFKKEEPMISTISVEAEWMACTQTELGHSVLSVAFSAD	852
OY	258	-ORIHCRS-----ETSKG-----VYLOY--DDOKIVSGLRLNTIKIMDK	294
DB	853	GGRVASGSDDKTIKIMDTASGTCGTQTELGHSQVSWVAAPSPREKRASSDDKTIKIMDA	912
OY	295	NILECKRIITGHTGSVLCQY--DERVLIITGSSDSTVRVWVNTGEMLTLIHCEAVLA	352
DB	913	ASGTCGTQTELGHSQVSWVAAPSPDQGVASGSDDHITIKIMDAASGTCGTQTELGHSQVSW	972
OY	353	LREN-NGMWYTC-SKRSTAVNMDMASPTDITLRLRVLGHRAAVNVYDF--DDKIYISAG	408
DB	973	VAPSPDQGVASGSDGKTIKIMDTASG---TCTQTELGHSQVSWVAAPSPDQGVASGSD	1029

Oy	409	DRTIKVNTSTCEPVRTLNGKRGLACLOQR--DRIVSSGSDNTITKLMEGCACLRVL	466
Oy	1030	DKTIKIMPTASTCTCTGLEGGGWQSVFSPDGORAVSGDDHTIKIMDAVSGTCTQTLL	1089
Oy	467	ECHHELTVLCIRF--DNKRIVSGAYDGKIKIWDLVAAALDPRAAGTCLRLTVHESGRYFR	524
Db	1090	EGHGGSWVSVAFPDGGQGVASGSIDGTIKIMD-----AASGT-CGTGLEGGWVHS	1140
Oy	525	LQF--DEFQIVSSSHDDTTLIMDFLNDPAQ	553
Db	1141	VAFSPDGORVASGSIDGTIKIMDAASGTCTQT	1171
RESULT	10		
T5136			
WD repeat protein popl [imported] - fission yeast (Schizosaccharomyces pombe)			
C:Species:	Schizosaccharomyces pombe		
C:Date:	21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000		
C:Accession:	T45136; T40157		
R:Komami, K.; Toda, T.			
submitted to the EMBL Data Library, September 1996			
A:Description:	Fission yeast WD repeat protein Popl is involved in maintenance of plo		
A:Reference number:	22925		
A:Accession:	T45136		
A:Status:	Preliminary; translated from GB/EMBL/DDBJ		
A:Molecule type:	DNA		
A:Residues:	1-775 <KOM>		
A:Cross-references:	EMBL:Y08391; PIDN:CAB69671.1		
A:Experimental source:	Strain h-972		
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Devlin, K.; Churcher, C.M.			
submitted to the EMBL Data Library, March 1998			
A:Reference number:	221842		
A:Accession:	T40157		
A:Status:	Preliminary; translated from GB/EMBL/DDBJ		
A:Molecule type:	DNA		
A:Residues:	1-775 <WOOD>		
A:Cross-references:	EMBL:AL022103; PIDN:CAI17898.1; GSPDB:GN00067; SPDB:SPBC262.18		
A:Experimental source:	Strain 97Zh-; cosmid c2g2		
C:Genetics:			
A:Gene:	SPBC262.18		
A:Map position:	2		
A>Note:	popl+		
Query Match	14.9%; Score 453; DB 2; Length 775;		
Best Local Similarity	25.8%; Pred. No. 4, 9e-27;		
Matches 162; Conservative	86; Mismatches 229; Indels 150; Gaps 20;		
Oy	9	QEAKLFPMNSSEBDC-----NNGEPRKKIIPKSNLSRQTY--NSCARCLINOET	56
Db	121	REKCLRNRSSLSNHLANKRFLFNSSODGMKNKEFFPTNINSVFPPNCKSKEYASET	180
Oy	57	VC-----LASTAMKTENCVAATKLTANGTSMIYPKOKRLSA-----	92
Db	181	TFELDLPNNNSVNVSYSPNLMLGDMSKTRQSPRHSSSSHNLSHEPVIVDFSENPSIH	240
Oy	93	----STEKEREL---CVKTFEQMESDOVEYEHLISOMCHIQGHINSLYLKPIORNF	144
Db	241	SNHLSSSKNVALELAOLISSFEKLPEPSVRQYLEFHILSLSCGRHAQONHKILLPLFOKNE	300
Oy	145	ITPLAPRGDIHIENITSYDAKSLCAEALVCEKYR--TSPGMLMKUL-----I	193
Db	301	LTEGPA-----ETINLVLIHDAPSLCAVSOVSHHWKLYVSNBELMKSFLKDGFPMDSI	356
Oy	194	ERNVFTDSLWRGLAEERGWQYLEFKNKPPDGAANPSFYBALPYKTIIDIETTESNWRCG	253
Db	357	DSKIRFMCCLOSLSA-----CAIMKRYYRFHFMLRERW---	389
Oy	254	RHSLQRIHCSETSKGYYC--LOYDDOKIVSGLRDNTIKIMDKNMLECKRILTCHTGVS	310
Db	390	LHAPEKTKRCSFPIHGRLITLKLOFDODKIIVYSTCSPRINIYDTKITGVILRSLEHEDGY	449

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-703 <JAL>
 A:Cross-references: EMBL:AF064515; NID:93293382; PIDN:AAC39496.1; PID:93293383
 C:Genetics:
 A:Gene: pop2; sucl; SPAC4D7.03
 A:Map position: 1
 C:Function:
 A:Description: required to prevent spontaneous re-replication

Query Match 13.1%; Score 396.5; DB 2; Length 703;
 Best Local Similarity 23.3%; Pred. No. 1e-22;
 Matches 146; Conservative 92; Mismatches 226; Indels 163; Gaps 25;

```

14 KFMNSERED--CNGEPPRK-----IPKNSLRQTVNSCARCL---NETVCL 59
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
96 KFMNSERED--CNGEPPRK-----IPKNSLRQTVNSCARCL---NETVCL 154
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 ASYAMTENCVAKTKLANCTSSMIVPKOKKLSASYEK-EKELCVKFEQMSQDVEEVE 118
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
155 NSISSNSDNPSPPKY--DTSNVSFGSKPISDEDLNLSQSVOTFEEDLPFGIQ--STAF 211
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
119 HLISQCHVQHG--INSYLPMLQRPDITLPAKGLDHAENILSYLDAKSLCAAEVCK 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
212 FQLLSCNROSRLNLECEPLKDIENLPE---SIVQSLILNDLHSLSCRLVSP 267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
178 EWRV---TSDGMLMKLIERNVRTDSIMRGLAERKGGYLFKNKPPDGAPP---N 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
268 TNNRILDVHTS---YKMHF-----SLFGQINEND---RYANENLRPPFLHND 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 SEYRALYPKIIDIEITIESNMWRCGRHSIOR----- 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
313 QISDDEFFEIFK-----RHFLNRKRWLPSPISPSHSPPIHVPNMTISLLH 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 -----HCRSETSK-GVYCLQYDDQIVSGLENDTIKIMDKNT 296
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 KKRITTSSTGTTQIHNAITGVLEARLEKGGKGVAAVKNHENTLVSGSIDKIVRANITEK 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
297 LECKRLTGHGTSVLCQY-----DERITIGSSDSRYRVWDVNTGEML 340
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
421 AACTHIFRHSIITICLDELIVSRILRHGVEIVEDDQPIYVSGSDHDLRW----- 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 NTLIHCEAVLHLRFNNGMMVTCSDRSIAVDMASPTDITLRLVGHRAAVNVDFD 400
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
473 -----KLPRKTPDPYLPDNTNSIDRME-KNPFV---HLLIGHTSVKRTISGY 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 KTVASASGRTIKVWNTSGCEVRLNHHKGIACLOTRD--RLVYSSSSDNTIRLMDIE 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
518 DILVSGSYDSSIRIRVSTGDELYLRHSHSLRIYSVLEPERNICISGSMKSIKRWDL 577
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 CGACLRVLEGHEELVRCIRFDPKRTVSGAYDGKIKVWDLVAADPRRAGTLCRTIVEH 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
578 TGCTKYVLEGHDAFYTLNVRONRLISSASDSTRIMW-----LMTGPSN 628
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
519 SGVRFRLQDFDEFQIVVSSHDDTILMD 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
629 SGYISSFVSDEHKIT-SCNDGSKVLMD 654
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13
 S36113
 LIS-1 protein - human
 C:Species: Homo sapiens (man)
 C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 26-May-2000
 C:Accession: S36113
 R:Reiner, O., Carozzo, R., Shen, Y., Wehnert, M., Faustiniella, F., Dobyns, W.B., Caskey
 Nature 364, 717-721, 1993
 A:Title: Isolation of a Miller-Dieker lissencephaly gene containing a G protein beta-subunit
 A:Reference number: S36113; MUID:93361119
 A:Accession: S36113
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-409 <RAS>
 C:Genetics:
 A:Gene: LIS-1
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:103-136/Domain: WD repeat homology <WD1>
 F:145-178/Domain: WD repeat homology <WD2>
 F:187-220/Domain: WD repeat homology <WD3>
 F:229-262/Domain: WD repeat homology <WD4>
 F:333-366/Domain: WD repeat homology <WD5>
 F:375-409/Domain: WD repeat homology <WD6>

Query Match 11.7%; Score 354; DB 2; Length 409;
 Best Local Similarity 29.2%; Pred. No. 9.3e-20;
 Matches 87; Conservative 57; Mismatches 110; Indels 44; Gaps 8;

```

280 IVSGLRDNTIKIMKNTLECKRLTGHGTSVLCQYDF--RVITGSSDSRYRVWDVNTG 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 MVSASEDNTIKVWDETDGFEETLKGHTDSVODISFHSGLKSLASCADMTIKLMDPQGF 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
338 EMLNTLIHCEAVLHLRF--NNGMMVTCSDRSIAVDMASPTDITLRLVGHRAAVNV 395
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 ECIRTMGHCHDNVSSVAIMPNGDHIVASRDKTIKMEVOTGYCV---KFTGHRVVRM 238
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
396 V--DEDDKIYISASGDRITIKVWNTSGCEVRLNHHKGIACLOTRD----- 441
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
239 VRPNQDGTLLASCNDQTVRVVAVATKCKAELREHEHVECTISMAPESSYSISEATGS 298
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
442 -----LVYSSSSDNTIRLMDIEGACRLVLEGHEELVRCIRFDPN--KRIVSGAYD 489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 ETKSGKRGPFLLSGSRKRTIKMDVSTGKMLTVLGHDMNVRKGVLFHSGKFLISCAD 358
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
490 KRIWMDLVAALDPRAPAGTLCRTIVEHSGRVFRLODFE--QIVSSSHDDTILMD 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
359 KTLRVWDY-----KMKRCKMTLANHEHEFTVSLDFHTAPYVTVGSVDQYVKWE 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
 S48052
 Platelet-activating factor acetylhydrolase 45k chain - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 26-May-2000
 C:Accession: S48052; S51606; S48075
 R:Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
 Nature 370, 216-218, 1994
 A:Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activat
 A:Reference number: S48052; MUID:94301405
 A:Accession: S48052
 A:Molecule type: mRNA
 A:Residues: 1-410 <HAT>
 A:Cross-references: EMBL:D30615
 A:Note: the title has been revised in S48075
 R:Hattori, M.; Adachi, H.; Tsujimoto, M.; Arai, H.; Inoue, K.
 Nature 370, 391-391, 1994
 A:Title: Miller-Dieker lissencephaly gene encodes a subunit of brain platelet-activat
 A:Reference number: S48075
 A:Contents: annotation
 A:Note: this is a revision to the title from reference S48052
 R:Hattori, M.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S51606
 A:Accession: S51606
 A:Molecule type: mRNA
 A:Residues: 1-25, 'A', 27-153, 'E', 155-410 <HAM>
 A:Cross-references: EMBL:D30615; NID:9516665; PIDN:BA06305.1; PID:9551
 C:Superfamily: unassigned WD repeat proteins; WD repeat homology
 F:104-137/Domain: WD repeat homology <WD1>
 F:146-179/Domain: WD repeat homology <WD2>
 F:188-221/Domain: WD repeat homology <WD3>
 F:230-263/Domain: WD repeat homology <WD4>
 F:289-325/Domain: WD repeat homology <WD5>
 F:334-367/Domain: WD repeat homology <WD6>
 F:376-409/Domain: WD repeat homology <WD7>

Query Match	11.7%;	Score 354;	DB 2;	Length 410;
Best Local Similarity	29.2%;	Pred. No. 9.4e-20;		
Matches	87;	Conservative	57;	Mismatches 110;
			Indels	44;
			Gaps	8;

```

QY 280 IVSGBRNTKIKMDKNLECKRILITGSGVLQDYE--RVIITSSDSSTVAWVUNTG 337
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 MWSASEDATIKWMEYETGDERPILKHTSDSVODISPHSKKLASCADMTIKLMDPQEG 182
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 338 EMLNTLLIHCEAVLHLP--NGMAYVCSKDRISAIVWDMASPTDITLRLVGHRAAVV 395
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 183 ECIPTMGHNDHNVSSVAIMPNGDHIVASBDSKTIKMEVOTGYCV--KTFTHREMVVR 239
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 V--PEDDKTIYVSGDGTIKVWNTSTCEFRITLNGHKGAGLCAYDR----- 441
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 VRPNQDGTLLASGSDNDQIVRWVYVATECAELREHNVVECTSMAPESSTSEATGS 299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 -----LVYSGSSDNTIWMIDIEGACLRVLEGEHELVRILKEDN--KRIYSGAYD 489
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 300 ETKSKSGRPFILGSGSDSKTIKMDVSTGACIMTVIHDMMVGVGLFHSKGKFTLSCAD 359
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 490 GKIVMDLVALLDRPAGTLCITLIVHSGRVRLQDFEE--QIVSSHDDIILMD 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 360 KTLRWMD-----KNRKQMTLANHEHVFISDFHTAPRVYGSVDQIVKWE 408
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

hypothetical protein YCR072c - yeast (*Saccharomyces cerevisiae*)
C:Species: *Saccharomyces cerevisiae*
C:Date: 31-Mar-1992 #sequence: cerevisiae
C:Revision: 04-Dec-1992 #text_change: 26-May-2000
C:Accession: S19487; S26657
R:Ballestra, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19486
A:Accession: S19487
A:Molecule type: DNA
A:Residues: 1-484, 'I', 486, 'TKL', 490, 492-515 <BAl1>
A:Cross-references: EMBL:X59720; MIPS:YCR072C
A:Note: this sequence has been revised in reference S26657
R:Ballestra, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, October 1992
A:Reference number: S26587
A:Accession: S26657
A:Molecule type: DNA
A:Residues: 481-503 <BAl2>
A:Cross-references: EMBL:X59720; MIPS:YCR072C
C:Genetics:
A:Map position: 3R
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
F:182-215/Domain: WD repeat homology <WD2>
F:225-265/Domain: WD repeat homology <WD3>
F:274-306/Domain: WD repeat homology <WD4>
F:398-431/Domain: WD repeat homology <WD5>
F:440-473/Domain: WD repeat homology <WD6>

[illegible][illegible]

Search completed: May 8, 2002, 10:51:11
Job time: 123 sec


```

FT      /note="WD motif"
FT      Region      516..544
FT      /note="WD motif"
XX      MO9938969-A1.
XX      PD
XX      05-AUG-1999.
XX      PF
XX      29-JAN-1999: 99MO-PR00196.
XX      PR
XX      09-DEC-1998: 98PR-0015545.
XX      30-JAN-1998: 98PR-0001100.
XX      PA
XX      (INRM ) INSEPM INST NAT SANTE & RECH MEDICALE.
XX      PI
XX      Arenzana Seisdedos F, Benarous R, Concordet J, Durand H;
XX      Kroll M, Margottin F;
XX      WPI: 1999-469329/39.
XX      DR
XX      N-PSDB; AAX86501.
XX      PT
XX      New human beta-transducin repeat containing protein and its
XX      fragments useful as, or to screen for, antiviral, antitumour,
XX      anti-inflammatory and anti-Alzheimer's agents
XX      PS
XX      Claim 1: Page 60-61; 71pp; French.
XX      CC
XX      The present sequence represents a human beta-transducin repeat containing
XX      protein (beta-Trcp). The protein directs proteins to the proteosome
XX      degradation pathways. The protein is able to interact with the Vpu
XX      protein of human immune deficiency virus-1 (HIV-1), cellular proteins
XX      Ikappab or beta-catenin (bc) and/or protein Skp1p. The protein controls
XX      ubiquitinylation of phosphorylated proteins and thus their targeting to
XX      proteosomes for degradation. Depending on whether the process is
XX      inhibited or promoted, the result may be delayed breakdown of CD4 (in
XX      cases of HIV-1 infection); increased activity of Ikb (and thus reduced
XX      activity of NFkappab) and increased degradation of mutant bc in tumour
XX      cells, or increased bc survival (and reduced apoptosis) in Alzheimer's
XX      patients. The beta-Trcp protein, and its active peptide fragments, or its
XX      nucleic acid, are used to screen for anti HIV-1 agents (antivirals),
XX      antitumour agents that disrupt cell cycle regulation or protein
XX      degradation in human tumour cells, and anti-inflammatory agents that
XX      disrupt activation by NFkappab. Fragments of the protein are also
XX      useful for treating osteo-articular inflammation or acute inflammation
XX      associated with release of tumour necrosis factor.
XX      SQ
XX      Sequence 569 AA:

```

```

Query Match      100.0%: Score 3034; DB 20; Length 569;
Best Local Similarity 100.0%: Pred. No. 3.2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 MDPAAVLOEKALKFMNSEREDCNGNPPRRITPEKNSLQNTVNSCARLCLNGETVLA 60
DB      1 mdpaavloekalkfmnserecngnpprritpeknsliqntvnscaarlclngetvla 60
OY      61 STAMKTCNCVAKTKLANGTSSNIVPKORKLSASYEKEKELCVKFEQSESDQVEFVHL 120
DB      61 stamktcncvaktklangtssniyvpkorklsasyekekelcvkfeyqsestdqvefhl 120
OY      121 ISQMKHYOHGHINSLAKMLORDPITPALPANGLDHIAENILSYLDANSLSAELVCKEWY 180
DB      121 isqmkhyohghinslylklpmlgrdfiltalpargldhiaenilsyldakslsaelyvckewy 180
OY      181 RTSQGMILMKKLIERRVATDSLMRGLARRGSGQYLFRNKKPPDGNAPNSRYALLYKTI 240
DB      181 rtsqgmilmkklierrvatdslmrglarrgsgqylfrnkkppdgnappnsryallyktti 240
OY      241 ODIETIESNMKGRHSIORICRSETSKGVYCLQYDDQKIVSGJGLRDMTIRKIMDNTLECK 300
DB      241 qdietiesnmkgrhsioricrsetskgyvclqyddqkivsgjglrdmtirkimdntleack 300

```

```

OY      301 RILTGHTGSVLQYDERVITITGSSDSTVRWVDYNTGEMLNTLIHCEAVLHLRFNNGM 360
DB      301 rilgtgtsvclqydevrititgssdstvrwvdyntgemlntlihcavlhrlfnngm 360
OY      361 VTCSKDRSIANWDMASPDITLRLRVLVGHRANVAVDPDDKYYIASGSDRTIKWNTSTC 420
DB      361 vtcskdrsiavwdmaspditlrlrvlvghranvavdpddkyyiasgsdrtikwnstsc 420
OY      421 EEFRTLNCHKRGKIGACLOVRDLVYSGSSDNTIRLMDIECGACLRVLEGHEELVRCIRFDN 480
DB      421 eefrtlnchkrgkiaclovrdlvyssgsdntirlmdiecgacrlrvlegheelvrcirfdn 480
OY      481 KRIVSGAYDGKIKVMDLVVALDPPRAGTCLRTLVESHSGRVRLQDFEQIYSSSHDT 540
DB      481 krivsgaydgkikvmdlvvaldppragtclrtlveshsgrvrlqdfeqiyssshdt 540
OY      541 ILIMDPLNDPAAQAEPPRSPRTYISR 569
DB      541 ilimdpndpaaqaepprspptyisr 569

```

```

RESULT 2
AAB12813
ID AAB12813 standard; protein; 569 AA.
XX
XX AAB12813:
XX
XX 27-NOV-2000 (first entry)
XX
XX

```

```

XX      Human beta-transducin repeat containing protein (beta-Trcp) SEQ ID NO:3.
XX      KW
XX      Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikappab;
XX      beta-catenin; Skp1; Cull1; F-box motif; WD40 repeat motif; FWD1;
XX      gene therapy; colon cancer; beta-transducin repeat containing protein;
XX      beta-Trcp.
XX      OS
XX      Homo sapiens.
XX      PN
XX      JP2000166542-A.
XX      PD
XX      20-JUN-2000.
XX      PF
XX      02-DEC-1998: 98JP-0343437.
XX      PR
XX      02-DEC-1998: 98JP-0343437.
XX      PA
XX      (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
XX      DR
XX      WPI: 2000-485550/43.
XX      N-PSDB; AAA73132.
XX
XX      F-box protein of ubiquitin ligase SCF complex which promotes the
XX      ubiquitination of Ikappab or beta-catenin
XX      PS
XX      Claim 3: Page 10-12; 19pp; Japanese.
XX
XX      The present invention describes an F-box motif protein of ubiquitin
XX      ligase SCF complex which promotes the ubiquitination of Ikappab or
XX      beta-catenin and is constituted by Skp1 protein, Cull1 protein and a
XX      complex (SCF complex) of F-box protein containing F-box motif and WD40
XX      repeat motif and has the amino acid sequence of 45 residues (AAB12811)
XX      or one of two 569 residue sequences (AAB12812, which is mouse ubiquitin
XX      ligase FWD1 protein) and (AAB12813, which is human beta-transducin
XX      repeat containing protein (beta-Trcp)). The F-box protein can be used for
XX      the gene therapy of colon cancer by being recombined to a virus vector.
XX      SQ
XX      Sequence 569 AA:

```

```

Query Match      100.0%: Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%: Pred. No. 3.2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MDPAAVLAQKALFEMNSEREDCNNGEPKRIPEKNSLRQTYNSCARCLNOETVCLA 60
DB 1 mdpaaavlaqkalkfemnserecdngeppkkipknslrqtynscaclngetvcia 60
QY 61 STAKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFEQWSESDQVEFVEHL 120
DB 61 staktekncvaktklantgssmivpkqkrlsaasyekekelcvkyfeqwsesdqvefvehl 120
QY 121 ISOMCHYOHGINSYLRKPMLOPDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWMY 180
DB 121 isomchyohghinsylrkpmldpftalparglldhiaenilsyldakslaaelvckewy 180
QY 181 RVTSDGMLMKRLIERMVRTDSLMRGLAERGMGOYLKKNPPDGNAPNSFYRALYPKII 240
DB 181 rvtsgdgmllwkkliermvrtdslmrglaermgoylknppdgnapnsfyralypkii 240
QY 241 QDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDOKIYSGLRDNTIKIMDKNTLECK 300
DB 241 qdietiesnwcgrhsioricrsetskgyvclqyddqkivsglrdntikimdkntleack 300
QY 301 RILTGHTGSVLCLQYDERVITITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 rilghtgsvlclqyderivititgssdstvrvwdvntgemlntlilhceavhlhfrfngmm 360
QY 361 VTCSKDRSIAVWDMASPDITLRRVLVGHRAAVNVDPDKIYASGDRITIKWNTSTC 420
DB 361 vtcskdrsiavwdmaspditlrrvlgvghraavnvdpdkiyasgdrityikwnstc 420
QY 421 EHVRTLNGHKGRIACLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
DB 421 ehvrtlnghkgriacloyrdrlvyssgndntirlmdiegacrlvleghheelvrcirfdn 480
QY 481 KRIVSGAYDGKIKVWDVLAALDPRAAPAGTLCRLTVLHSGRVFLQDFEQIYSSSHDT 540
DB 481 krivsgaydgkikvwdlvaaaldrapagtlcrltlvehsgrvflqdfefqivssshdt 540
QY 541 IILWDFLNDPAAQAEPPSPSRITYISR 569
DB 541 iilwdflnppaaqaeppsrityisr 569

RESULT 3
AA96697
ID AA96697 standard; Protein; 569 AA.
AC AA96697;
XX
XX 26-SEP-2000 (first entry)
DT
XX
XX Human beta-TrCP.
DE
XX E3 ubiquitin ligase; beta-TrCP; F-box; WD protein; I-kappa-B; inhibitor;
KW nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
KM anti-inflammatory; immunosuppressive; cytostatic.
XX
XX Homo sapiens.
OS
XX
XX WO200034447-A2.
PN
XX 15-JUN-2000
PD
XX 10-DEC-1999; 99WO-US293771.
PF
XX 10-DEC-1998; 98US-0210060.
PR
XX (SIGN-) SIGNAL PHARM. INC.
PA (YISS) YISSUM RES & DEV CO.
XX
XX Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
PI Laron I, Yaron A;
XX
XX MPI; 2000-431294/37.

```

```

DR N-PDB; AAA51229.
XX
XX Polypeptide enhancing phosphorylated Ikbapab ubiquitination useful for
PT treating disorder associated with NF-kappaB activation e.g. cancer,
PT comprising amino acid sequence of human E3 ubiquitin ligase or its
PT variant
XX
XX Claim 21: Page 72-74; 77pp; English.
PS
XX
XX Human beta-TrCP, an F-box/WD protein family member, has been shown to
CC have homology to human E3 ubiquitin ligase (E3). E3 enhances
CC ubiquitination of phosphorylated I-kappa-B, an inhibitor protein of
CC nuclear factor kappa-B (NF-kappa-B). Understanding I-kappa-B
CC degradation via the ubiquitin pathway is useful for identifying
CC modulators of this process for use in treating diseases associated with
CC activation of NF-kappa-B. In vitro analysis suggests that deletion of
CC the F-box results in a protein that functions as a dominant negative
CC molecule in vivo. Transient over-expression of delta-beta-TrCP (a
CC deletion mutant) inhibited the degradation of endogenous I-kappa-B-alpha
CC in stimulated Jurkat cells, resulting in accumulation of phosphorylated
CC I-kappa-B-alpha. E3 can be used to screen for modulators of NF-kappa-B
CC activity. E3 and beta-TrCP can be used to modulate NF-kappa-B to treat
CC inflammatory diseases, autoimmune diseases, cancer and viral infections.
XX
XX Sequence 569 AA:
SQ
Query Match 100.0%; Score 3034; DB 21; Length 569;
Best Local Similarity 100.0%; Pred. No. 3, 2e-287;
Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPAAVLAQKALFEMNSEREDCNNGEPKRIPEKNSLRQTYNSCARCLNOETVCLA 60
DB 1 mdpaaavlaqkalkfemnserecdngeppkkipknslrqtynscaclngetvcia 60
QY 61 STAKTEKNCVAKTKLANGTSSMIVPKOKRLSASYEKEKELCVKFEQWSESDQVEFVEHL 120
DB 61 staktekncvaktklantgssmivpkqkrlsaasyekekelcvkyfeqwsesdqvefvehl 120
QY 121 ISOMCHYOHGINSYLRKPMLOPDFITALPARGLDHIAENILSYLDAKSLCAELVCKEWMY 180
DB 121 isomchyohghinsylrkpmldpftalparglldhiaenilsyldakslaaelvckewy 180
QY 181 RVTSDGMLMKRLIERMVRTDSLMRGLAERGMGOYLKKNPPDGNAPNSFYRALYPKII 240
DB 181 rvtsgdgmllwkkliermvrtdslmrglaermgoylknppdgnapnsfyralypkii 240
QY 241 QDIETIESNMRCGRHSIORICRSETSKGVYCLQYDDOKIYSGLRDNTIKIMDKNTLECK 300
DB 241 qdietiesnwcgrhsioricrsetskgyvclqyddqkivsglrdntikimdkntleack 300
QY 301 RILTGHTGSVLCLQYDERVITITGSSDSTVRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
DB 301 rilghtgsvlclqyderivititgssdstvrvwdvntgemlntlilhceavhlhfrfngmm 360
QY 361 VTCSKDRSIAVWDMASPDITLRRVLVGHRAAVNVDPDKIYASGDRITIKWNTSTC 420
DB 361 vtcskdrsiavwdmaspditlrrvlgvghraavnvdpdkiyasgdrityikwnstc 420
QY 421 EHVRTLNGHKGRIACLOYRDRLVYSGSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN 480
DB 421 ehvrtlnghkgriacloyrdrlvyssgndntirlmdiegacrlvleghheelvrcirfdn 480
QY 481 KRIVSGAYDGKIKVWDVLAALDPRAAPAGTLCRLTVLHSGRVFLQDFEQIYSSSHDT 540
DB 481 krivsgaydgkikvwdlvaaaldrapagtlcrltlvehsgrvflqdfefqivssshdt 540
QY 541 IILWDFLNDPAAQAEPPSPSRITYISR 569
DB 541 iilwdflnppaaqaeppsrityisr 569

RESULT 4

```


AAV83041
 ID AAV83041 standard; Protein; 569 AA.
 AC AAV83041;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein FBP-1.
 XX
 KW F-box protein; FBP; diagnosis; treatment; screening; agonist;
 KW antagonist; proliferative disorder; differentiative disorder;
 KW breast cancer; prostate cancer; ovarian cancer; cancer;
 KW small cell lung carcinoma; immune disorder; cardiovascular disorder;
 KW inflammatory disorder; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200012679-A1.
 PD 09-MAR-2000.
 XX
 PF 27-AUG-1999; 99WO-US19560.
 XX
 PR 28-AUG-1998; 98US-0098355.
 PR 03-FEB-1999; 98US-0118568.
 PR 15-MAR-1999; 99US-0124449.
 XX
 PA (UYNY) UNITV NEW YORK STATE.
 XX
 PI Chlaur DS, Fagano M, Latres E;
 XX
 DR WPI: 2000-256635/22.
 DR N-PSDB; AA293350.
 XX
 PT Novel nucleic acid for screening compounds useful for treating
 PT proliferative and differentiative disorders such as cancer and immune
 PT disorders comprises sequences encoding ubiquitin ligases
 XX
 PS Disclosure; Figure 3a; 245pp; English.
 XX
 CC Nucleic acids encoding substrate-targeting subunits of ubiquitin
 CC ligases with F-box motifs (F-box proteins) are useful for diagnosis
 CC of proliferative and differentiated related disorders by measuring
 CC FBP gene expression. Cells expressing such proteins or
 CC their fragments are useful for screening compounds. The compounds
 CC are agonists or antagonists, which are useful for treating a
 CC proliferative or differentiative disorder in a mammal such as
 CC breast, ovarian and prostate cancer and small cell lung carcinoma
 CC and also major opportunistic infections, immune disorders,
 CC cardiovascular diseases and inflammatory disorders. FBP protein,
 CC analogs, derivatives and their subsequences, anti-FBP antibodies
 CC are also useful in diagnosis of the disorders.
 XX
 SQ Sequence 569 AA;

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 3.2e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPPEAVYQKALFKMNSEREDNNGEPRIKIIPEKNSLRQYNSCARCLNOETVCLA 60
 DB 1 mdppeavylqekalkfmsserednngpeprkiipeknslrqynscarcclnoetvcla 60
 QY 61 STAMKTENCYAKTKLANGTSSMIVPKQKLSASYEKELCVYFEQWSESDVEFEHL 120
 DB 61 stamktenccvaktklangtssmivpkqklsasyekelcvkyfegwssdqvfehl 120
 QY 121 ISOMCHQHOHINSYKPMQROFITLAPARGLDHTAENLSTLDKSLCAALVCWEKY 180
 DB 121 isgmchqghnhsylkpmldrftitlpargldhtaenllstldakslcaaelvcweky 180
 QY 181 RVTSQGMILMKRLIERMVRTSLMRLAERGMQYLFKNKPPGNAFPNSFYALPKIT 240

DB 181 rvtsqgmilwkkllermvrtslwrglaerrwgqylfknpkgnaapnsfyalypkii 240
 QY 241 QDIETIESNWRGSRHSIORICHKSETSKGVYCLQYDQKIVSGLRDNTIKIMDKNTLECK 300
 DB 241 qdietiesnwrgrhsqirhcrseetskgyclqyddqkivsglrdntlikwknleek 300
 QY 301 RILGHTGSVLCIYDERVITITGSSDSYVWVDVNTGEMLNTLIHCEAVLHIRENNGM 360
 DB 301 rilghtgsylciqyderviltgssdstyvwvntgemlnllihceavllhfrnngm 360
 QY 361 VTCSSKDSIIVWMASTPTDTERVLYGHRAAVNVDPDKYIVSASGRITKVMWTSNC 420
 DB 361 vtcsskdsiivwmasptdlttrvlyghraavnvdpdkyivsagrdtkvmwtstc 420
 QY 421 EFWRTLNGHRRGIACLOYRDLVYSSSDNTIRLMDIEGACALRYEGHEELVRCIRPDN 480
 DB 421 efwrtlnghkrgiacloyrdrlvysssdntlrlmdiegcacrlryegheelvrclrfdn 480
 QY 481 KRIVSGAYDGKIKVMDLVVALDPRAPAGTLCRTLYEHSGRVRLQDFEQIVSSSDHT 540
 DB 481 krivsgaydgkikvmdlvvaaloprpagtclrtlyehsgrvrlqdfeqivssshdt 540
 QY 541 ILIWFELNDPAAQAEPPRSPSRITYYISR 569
 DB 541 iliwfndpaaqaepprsprityylsr 569

RESULT 5
 AAT83250
 ID AAV83250 standard; Protein; 569 AA.
 AC AAV83250;
 XX
 DT 16-AUG-2000 (first entry)
 XX
 DE F-box protein hBetatrCp.
 XX
 KW Ubiquitin ligase; SCF; F-box protein; targeted degradation;
 KW destabilisation; proteolysis; drug discovery; gene therapy; cancer;
 KW oncoprotein; Huntington's disease; gene knockout; delivery systems;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN WO200022110-A2.
 XX
 PD 20-APR-2000.
 XX
 PR 08-OCT-1999; 99WO-US23705.
 PR 09-OCT-1998; 98US-0103787.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Zhou P, Howley P;
 XX
 DR WPI: 2000-317970/27.
 DR N-PSDB; AA293710.
 XX
 PT Targeting degradation of polypeptide useful for treating cancer and
 PT other proliferative disorders, involves conjugating polypeptide with
 PT ubiquitin protein ligase or inhibiting ubiquitination using organic
 PT compound
 XX
 PS Claim 9; Page 171; 185pp; English.

The F-box proteins are a family of ubiquitin ligases (SCF ubiquitin
 ligases) which can be used for the targeted degradation of a target
 polypeptide in vivo. Targeted degradation is achieved by expressing
 the ubiquitin ligase in a cell linked to the interaction domain of
 the target polypeptide and thereby recruiting the target polypeptide
 to the ubiquitin ligase. Such methods are useful for decreasing or

xx Human cell signaling proteins useful for, e.g. diagnosing cell
 pt proliferative and inflammatory disorders -
 ps Claim 1; Page 77-78; 90pp; English.

CC The present sequence is cell signalling protein-12 (CSIGP-12) encoded
 CC by cDNA obtained from Inocyte clone 3239149 of COLACCTOL library. It is
 CC expressed in musculoskeletal, gastrointestinal and nervous tissues and is
 CC found to be homologous to beta-transducin repeats containing
 CC protein. Fragments of CSIGP encoding nucleic acid can be used as
 CC hybridisation probe for detecting CSIGP related sequences or allelic
 CC variants. Recombinant CSIGP can be produced in host cells by transforming
 CC them with genetically engineered vectors. Agonists or antagonists can be
 CC used in the treatment of cell proliferative and inflammatory disorders
 CC associated with decreased or increased CSIGP expression. CSIGP is used in
 CC the diagnosis, prevention and treatment of cell proliferative disorders
 CC like arteriosclerosis, cirrhosis, cancer, hepatitis and inflammatory
 CC disorders like AIDS, Addison's disease, multiple sclerosis, etc.

xx Sequence 569 AA:

Query Match 100.0%; Score 3034; DB 21; Length 569;
 Best Local Similarity 100.0%; Pred. No. 3.2e-287;
 Matches 569; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNNSSREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 mdpaeavlgekalikfnnsseredcngpeprkllpeknslrqltynscarcilnqetvcla 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORKLSASYEKEKELCVYFPQWSESDQVEVEHL 120
 DB 61 stamkencvaktklngssmtvvpkorklsasyekekelcvyfpqwsesdqvetvehl 120
 QY 121 ISOMCHYOHGHINSYKLPMLQDRDFTALPARGLDHAENTLSYLDKSLCAELVCKEWY 180
 DB 121 isgmchyghghinsylkplmqdrdftalpargldhaentlsyldakslcaaelvckewy 180
 QY 181 RYTSDDMLMKKLLERVRRTDSLRGLAERGWQCYLFKNKPPGNNAPNSFYALPKIT 240
 DB 181 rvtsgdmlwkllewrtrtdslwrglaerwqgylfknppgngnappnsfyalypkll 240
 QY 241 ODITTESNMRCGRHSLOIRHCRSETSKGYCYQYDQKIVSGLRDNTIKIMDKNLECK 300
 DB 241 qditlesnmrcgrhsloirhcrsetskgyccyqyddqkivsglrdntlikimdknleack 300
 QY 301 RIITGHTGSVLCQYDERVITIGSSDSTYRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 riltghtgsvlclqyderivitigssdstyrvwdvntgemlntlihnceavlhrlfningmm 360
 QY 361 VYCSKRSIAVWMDASPTDTLRRVLVGHRAAVNVYDFDQKYIVASGSDTITVWNTSTC 420
 DB 361 vycskrsiavwmdasptdtlrrvlvghraavnvdydfdqkyivassgdtkllywnsttc 420
 QY 421 EFVRTNGHKGRTAGLQYRDRLVSSSDNTIRLMDIEGACLRVLEGHELVRCIRFDN 480
 DB 421 efvrtngkhgrtaglqyrdrlvssssdntirlmdiegacrlrvleghelevrcirfdn 480
 QY 481 KRIVSGAVDGKIKVMDLVLAALDPRAPAGTCLRTLYVHSGRVRRLQDFEQIYSSSHDT 540
 DB 481 krtvsgavdgkikvmdlvlaaldrpapagtlclrtlyvhsgrvrirlqdfeqiyssshdt 540
 QY 541 ILIWDPLNDPAAQAEPRSPRTYTYISR 569
 DB 541 ililwdplndpaaqaeprrsprtytyistr 569

RESULT 7
 AAB48298
 ID AAB48298 standard; protein; 569 AA.
 XX
 AC AAB48298;

xx 02-APR-2001 (first entry)
 dt
 xx Human ZF11 protein.
 DE

xx S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 KW Bad; Bcl-2; tumour; cytostatic.
 xx

OS Homo sapiens.

PN WO200075184-A1.

PD 14-DEC-2000.

PF 05-JUN-2000; 2000MO-US15449.

PR 04-JUN-1999; 99US-0137494.

PA (UYA) UNIV YALE.

PI Zhang H, Tsvetkov LM, Kondo T;

DR WPI: 2001-061703/07.

DR N-PSDB; AAC84610.

PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
 pt involves altering levels of proteins such as S-phase kinase associated
 proteins 1, 2 and cullin/CDC53 proteins -

PS Claim 3; Page 130-132; 162pp; English.

xx The invention relates to methods of altering the polypeptide levels in a
 CC cell, using proteins selected from S-phase kinase associated proteins 1
 CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
 CC cullin/CDC53 family of proteins). The method is useful for altering the
 CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
 CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
 CC detecting tumors, and in monitoring tumor treatment in a mammal. Agents
 CC that modulate interactions between SKP and target proteins are useful for
 CC treating tumors.

xx Sequence 569 AA:

Query Match 99.8%; Score 3027; DB 22; Length 569;
 Best Local Similarity 99.8%; Pred. No. 1.6e-286;
 Matches 568; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPAEAVLOEKALKFNNSSREDCNNGEPKRIIPKNSLRQTYNSCARCLNOETVCLA 60
 DB 1 mdpaeavlgekalikfnnsseredcngpeprkllpeknslrqltynscarcilnqetvcla 60
 QY 61 STAMKENCVAKTKLANGSSMTVPKORKLSASYEKEKELCVYFPQWSESDQVEVEHL 120
 DB 61 stamkencvaktklngssmtvvpkorklsasyekekelcvyfpqwsesdqvetvehl 120
 QY 121 ISOMCHYOHGHINSYKLPMLQDRDFTALPARGLDHAENTLSYLDKSLCAELVCKEWY 180
 DB 121 isgmchyghghinsylkplmqdrdftalpargldhaentlsyldakslcaaelvckewy 180
 QY 181 RYTSDDMLMKKLLERVRRTDSLRGLAERGWQCYLFKNKPPGNNAPNSFYALPKIT 240
 DB 181 rvtsgdmlwkllewrtrtdslwrglaerwqgylfknppgngnappnsfyalypkll 240
 QY 241 ODITTESNMRCGRHSLOIRHCRSETSKGYCYQYDQKIVSGLRDNTIKIMDKNLECK 300
 DB 241 qditlesnmrcgrhsloirhcrsetskgyccyqyddqkivsglrdntlikimdknleack 300
 QY 301 RIITGHTGSVLCQYDERVITIGSSDSTYRVWDVNTGEMLNTLIHCEAVLHLRFNNGMM 360
 DB 301 riltghtgsvlclqyderivitigssdstyrvwdvntgemlntlihnceavlhrlfningmm 360

QY 361 VTCSKDRSIAVWDASPDITLRLVGVHRAVNVDPDDKTVASGSDRTIKWNTSTC 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 vtcskdrsiavwdaspditlrlvghraavnvdddktyvasgsdrtikwnstlc 420
 QY 421 EFVRLNGHKRGIACTLOYRDLVWSSSDNTIRLMDIEGACLRVLECHELVRCIRFDN 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 efvrlngkhrgiaclgyrdrlvsgssdntlrlwdiecgacrlvleghevlrcirfdn 480
 QY 481 KRIVSGAYDGKIKVWDIVAALDPRAPACTLCRLTVHSGRVFRLOPDEFOIVSSSHDT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 krivsgaydgkikvwdlvaaldrpapagtlcrltlvehsgvfrlqfdeifqvssshdt 540
 QY 541 ILIMFLNDPAAQAEPPSPSRTTYISR 569
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 ilimflndpaaqaepprsprtyltsr 569
 RESULT 8
 AAM00960
 ID AAM00960 standard; Protein: 608 AA.
 XX
 AC AAM00960;
 XX
 DT 01-OCT-2001 (first entry)
 DE Human bone marrow protein, SEQ ID NO: 436.
 XX
 KW Human: bone marrow; antiinflammatory; cytostatic; neuroprotective;
 KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
 KW immunosuppressive; gene therapy; cytokine cell proliferation;
 KW cell differentiation modulator; immune disorder; infection; cancer;
 KW human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153453-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 23-DEC-2000; 2000WO-US34960.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 30-NOV-2000; 2000US-0250583.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Ford JE, Boyle BJ, Tang YT, Liu C, Asundi V, Chen R, Ma Y;
 PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Drmanac RT;
 XX
 DR MPI; 2001-488707/53.
 DR N-PSDB; AAH90079.
 XX
 PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
 PT for treating e.g. cancer and immune deficiency disorders -
 XX
 XX Claim 10; Page 523-524; 648bp; English.
 CC The present sequence is one of 251 novel human polypeptides encoded
 CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
 CC polypeptide encoded by it are useful in the treatment of various
 CC immune deficiencies and disorders. The deficiencies and disorders may
 CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
 CC infection, or may result from an autoimmune disorder, a coagulation
 CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
 CC suppression of an inflammatory response or treatment of a nervous
 CC system disorder such as Alzheimer's disease. Detection of the presence

CC or increased expression of the polynucleotide or the protein it
 CC encodes is useful for the diagnosis and/or prognosis of one
 CC or more types of cancer. The polynucleotide and polypeptide can be
 CC used as nutritional sources or supplements and in the screening of
 CC chemical compounds as potential drugs.
 XX
 SO Sequence 608 AA;
 Query Match 99.8%; Score 3027; DB 22; Length 608;
 Best Local Similarity 99.6%; Pred. No. 1,7e-286;
 Matches 567; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDPAAVLQEKALKFNMSSEREDCNNGEPKKIIPKXNSLRQTYNSCARLCLADQEVCLA 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 40 mdpaaavlqekalkfnmsseredcngpeppkiklpeknslrqtnscarlclqetvcla 99
 QY 61 STAMKTEHCVAKTLANGTSMIYVKKORLSASYEKEKELCVKFEQWSDQVEFENL 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 100 stamktehcvaaktlangtsmiypvkkqlsaasyekelcvkfeqwsdqvefvehl 159
 QY 121 ISOMCHYOHGINSYLYKPMLODFITLPAKGLDHAENILSYLDAKSLCAELVCKEWY 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 160 isgmchyghinsylkpmldqdfitalparglidhaenilslsyldakslcaaelvckewy 219
 QY 181 RYTSQGLMKRLIERAVRTDSLMRGLAERKMGQYLKRNKPPDGNAPNSFYRLALPKII 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 220 rlysdgmlwkklieravrtdslwrglaerwggylkfnppdgnapnsfyalpkil 279
 QY 241 ODIEFIESNMWCGRHSIORICRSETSKGVYCLQYDOOKIVSGLRDWTIKIMDKNTLECK 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 280 qdiefiesnmwcrhsioricrsetskgyyclqyddkivsglrdwtiklwdkntleck 339
 QY 301 RILTGHTSVLCLOYDERVITIGSSDSTVRWVDVNTGEMTLIHCEAVLHREFNNGM 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 340 rilghtsvlclgydervitlgsdsstvrwvntgemltlilheavvlhrfnngm 399
 QY 361 VTCSKDRSIAVWDASPDITLRLVGVHRAVNVDPDDKTVASGSDRTIKWNTSTC 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 400 vtcskdrsiavwdaspditlrlvghraavnvdddktyvasgsdrtikwnstlc 459
 QY 421 EFVRLNGHKRGIACTLOYRDLVWSSSDNTIRLMDIEGACLRVLECHELVRCIRFDN 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 460 efvrlngkhrgiaclgyrdrlvsgssdntlrlwdiecgacrlvleghevlrcirfdn 519
 QY 481 KRIVSGAYDGKIKVWDIVAALDPRAPACTLCRLTVHSGRVFRLOPDEFOIVSSSHDT 540
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 520 krivsgaydgkikvwdlvaaldrpapagtlcrltlvehsgvfrlqfdeifqvssshdt 579
 QY 541 ILIMFLNDPAAQAEPPSPSRTTYISR 569
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 580 ilimflndpaaqaepprsprtyltsr 608
 RESULT 9
 AAB12812
 ID AAB12812 standard; Protein: 569 AA.
 XX
 AC AAB12812;
 XX
 DT 27-NOV-2000 (first entry)
 DE Mouse ubiquitin ligase FWD1 protein SEQ ID NO:2.
 XX
 KW Ubiquitin ligase SCF complex; F-box protein; ubiquitination; Ikapab;
 KW beta-catenin; Skp1; Cull; F-box motif; WD40 repeat motif; FWD1;
 KW gene therapy; colon cancer; beta-transducin repeat containing protein;
 KW beta-TrCP.
 XX
 OS Mus musculus.
 XX
 PN JP200016542-A.
 XX


```

QY 121 ISOMCHYGHINSYKPMLODFITLALPARGLDHAENIISYDAKSICAAELVCKEY 180
XX
DB 121 isqmchyghinsy kpmldgfdltalpargl dhaenil sydaksicaaelvckey 180
QY 181 RYTSOSGLMKKTIERNVRDLSMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKII 240
DB 181 r y t s o s g l m k k t i e r n v r d l s m r g l a e r g m g o y l f r k k p p d g n a p p s f y r a l y p k i i 240
QY 241 ODIEETESNMWRCGRHSIQRHCRSETSKGVCLQYDDOKIVSGLRNDNTIKIWDKNTLECK 300
DB 241 o d i e e t e s n m w r c g r h s i q r h c r s e t s k g v c l q y d d o k i v s g l r n d n t i k i w d k n t l e c k 300
QY 241 qdletiesnmwrcgrhs iqrhcrse tskgvclqyddokivsglrdntik iwdkntle ck 300
QY 301 RIITGTSVCLQYDERIYITGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGM 360
DB 301 r i i t g t s v c l q y d e r i y i t g s s d s t y r v w d v n t g e m l n l i h c e a v l h r e n n n g m 360
QY 361 VCSKSRSLAVWDMASPTDITLRVYGHRAAVNVDPDRTIVSASGRTIKWNTSTNC 420
DB 361 v c s k s r s l a v w d m a s p t d i t l r v y g h r a a v n v d p d r t i v s a s g r t i k w n t s t n c 420
QY 421 EFVRLNGHKGRIACLOYRDLRVSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDN 480
DB 421 e f v r l n g h k g r i a c l o y r d l r v s s g s d n t i r l m d i e c g a c l r v l e g e e l v r c i r f d n 480
QY 481 KRIVSGAYGKTKVMDLVALDPRAPAGTLCRLTVEHSGRFRLOFDFQIVSSSHDT 540
DB 481 k r i v s g a y g k t k v m d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d f q i v s s h d t 540
QY 541 ILIMDFLNDPAQAEPSPSRRTYYSR 569
DB 541 i l i m d f l n d p a a q a e p s p s r r t y y i s r 569

```

RESULT 11

AA00847
ID AA00847 standard; Protein: 590 AA.

XX
AC AA00847;

XX
DT 01-OCT-2001 (first entry)

XX
DE Human bone marrow protein, SEQ ID NO: 210.

KW Human: bone marrow; antiinflammatory; cyrostatic; neuroprotective;

KW antiviral; antibacterial; antifungal; anti-HIV; haemostatic;

KW immunosuppressive; gene therapy; cytokine cell proliferation;

KW cell differentiation modulator; immune disorder; infection; cancer;

XX human immunodeficiency virus; HIV; autoimmune disorder; haemophilia.

OS Homo sapiens.

XX
PN WO200153453-A2.

XX
PD 26-JUL-2001.

XX
PF 23-DEC-2000; 2000WO-US34960.

XX
PR 21-JAN-2000; 2000US-0488725.

XX
PR 25-APR-2000; 2000US-0552317.

XX
PR 09-JUL-2000; 2000US-0598042.

XX
PR 19-JUL-2000; 2000US-0620312.

XX
PR 03-AUG-2000; 2000US-0653450.

XX
PR 14-SEP-2000; 2000US-0662191.

XX
PR 19-OCT-2000; 2000US-0693036.

XX
PR 30-NOV-2000; 2000US-0250583.

XX
PA (HYSE-) HISEQ INC.

PI Ford JE, Boyle BJ, Tang YF, Liu C, Asundi V, Chen R, Ma Y;

PI Ren F, Wang J, Werhman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Drmanac RT;

XX
DR WPI; 2001-488707/53.

```

DR N-PSDB: AAH89966.
XX
PT Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT for treating e.g. cancer and immune deficiency disorders
XX
PS Claim 10; Page 354-355; 648pp; English.
XX
CC The present sequence is one of 251 novel human polypeptides encoded
CC by a bone marrow-expressed polynucleotide. The polynucleotide and the
CC polypeptide encoded by it are useful in the treatment of various
CC immune deficiencies and disorders. The deficiencies and disorders may
CC be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC infection, or may result from an autoimmune disorder, a coagulation
CC disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC suppression of an inflammatory response or treatment of a nervous
CC system disorder such as Alzheimer's disease. Detection of the presence
CC or increased expression of the polynucleotide or the protein it
CC encodes is useful for the diagnosis and/or prognosis of one
CC or more types of cancer. The polynucleotide and polypeptide can be
CC used as nutritional sources or supplements and in the screening of
CC chemical compounds as potential drugs.
XX
SQ Sequence 590 AA:

```

Query Match 97.68; Score 2962; DB 22; Length 590;
Best Local Similarity 99.88; Pred. No. 3.7e-280;
Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 15 FMNSEREDCNNGEPPRKTIPEKNSLRQTYNSCARLCLNCFVCLASTAMKTECVAKTK 74
DB 15 f m n s e r e d c n n g e p p r k t i p e k n s l r q t y n s c a r l c l n c f v c l a s t a m k t e c v a k t k 74
QY 36 fgnsseredcmngepprki t i p e k n s l r q t y n s c a r l c l n g e t v c l a s t a m k t e c v a k t k 95
DB 36 f g n s s e r e d c m n g e p p r k i t i p e k n s l r q t y n s c a r l c l n g e t v c l a s t a m k t e c v a k t k 95
QY 75 LANGTSSMIVPKQRKLSAYEKEKELCVKRFQWSESDQVEFVHLLISOMCHYGHINS 134
DB 75 l a n g t s s m i v p k q r k l s a y e k e k e l c v k r f q w s e s d q v e f v h l l i s o m c h y g h i n s 134
QY 96 langtssmivpkqrkl s a y e k e k e l c v k f e q s e s d q v e f v h l l i s q m c h y g h i n s 155
DB 96 l a n g t s s m i v p k q r k l s a y e k e k e l c v k f e q s e s d q v e f v h l l i s q m c h y g h i n s 155
QY 135 YLKPMLODFITLALPARGLDHAENIISYDAKSICAAELVCKEYRYSQGLMKKIIIE 194
DB 135 y l k p m l o d f i t l a l p a r g l d h a e n i l s y d a k s i c a a e l v c k e y r y s q g l m k k i i e 194
QY 156 ylkpmldgfdltalpargl dhaenil sydaksicaaelvckeyr ytsqglmkk i i e 215
DB 156 y l k p m l d g f d l t a l p a r g l d h a e n i l s y d a k s i c a a e l v c k e y r y t s q g l m k k i i e 215
QY 195 RMVRTDSLMRGLAERGMGOYLFRKKPPDGNAPPSFYRALYPKIIODIETIESNMWRCGR 254
DB 195 r m v r t d s l m r g l a e r g m g o y l f r k k p p d g n a p p s f y r a l y p k i i o d i e t i e s n m w r c g r 254
QY 216 rmvrtdslmrgrlaer g m g o y l f r k k p p d g n a p p s f y r a l y p k i i d i e t i e s n w r c g r 275
DB 216 r m v r t d s l m r g l a e r r g y g y l f r k k p p d g n a p p s f y r a l y p k i i d i e t i e s n w r c g r 275
QY 225 HSLQRIHCRSETSKGVCLQYDDOKIVSGLRNDNTIKIWDKNTLECKRIITGTSVCLQ 314
DB 225 h s l q r i h c r s e t s k g v c l q y d d o k i v s g l r n d n t i k i w d k n t l e c k r i i t g t s v c l q 314
QY 276 hslqrhcrse tskgvclqyddokivsglrdntik iwdkntle ckriitgtsvclq 335
DB 276 h s l q r i h c r s e t s k g v c l q y d d q k i v s g l r d n t i k i w d k n t l e c k r i i t g t s v c l q 335
QY 315 YDERYITGSSDSTYRVWDVNTGEMLNLIHCEAVLHRENNNGMVTCSKDRSLAVWDM 374
DB 315 y d e r y i t g s s d s t y r v w d v n t g e m l n l i h c e a v l h r e n n n g m v t c s k d r s l a v w d m 374
QY 336 yderiyitgssdstyrvwdvntgemlnli hceavlh l r n n g m v t c s k d r s l a v w d m 395
DB 336 y d e r i y i t g s s d s t y r v w d v n t g e m l n l i h c e a v l h l r n n g m v t c s k d r s l a v w d m 395
QY 375 ASPFDITLRVYGHRAAVNVDPDRTIVSASGRTIKWNTSTCEFRVRLNGHKGRIA 434
DB 375 a s p f d i t l r v y g h r a a v n v d p d r t i v s a s g r t i k w n t s t c e f r v r l n g h k r g i a 434
QY 396 aspfditlrrv yghraavnvdpdrtiv s asgdr t i k w n t s t c e f r l n g h k g r i a 455
DB 396 a s p f d i t l r r v y g h r a a v n v d d i d k y i v s a s g d r t i k w n t s t c e f r l n g h k g r i a 455
QY 435 CLQYRDLRVSSGSDNTIRLMDIECGACLRVLEGEELVRCIRFDNKRIVSGAYGKIV 494
DB 435 c l o y r d l r v s s g s d n t i r l m d i e c g a c l r v l e g e e l v r c i r f d n k r i v s g a y g k i v 494
QY 456 clqydrdrvssgsdntir lwdiecgac lrvle g e e l v r c i r f d n k r i v s g a y g k i v 515
DB 456 c l q y d r l v s g s d n t i r l w d i e c g a c l r v l e g e e l v r c i r f d n k r i v s g a y g k i v 515
QY 495 WDLVALDPRAPAGTLCRLTVEHSGRFRLOFDFQIVSSSHDTILIMFNDPAAQA 554
DB 495 w d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d f q i v s s s h d t i l i m f n d p a a q a 554
QY 516 wdlvaldprapag t l c r l t v e h s g r f r l o f d f q i v s s s h d t i l i w d f n d p a a q a 575
DB 516 w d l v a l d p r a p a g t l c r l t v e h s g r f r l o f d f q i v s s s h d t i l i w d f n d p a a q a 575
QY 555 EPPRSRSPRTYYSR 569
DB 555 e p p r s r s p r t y y i s r 569

```

RESULT 12

AA085852
ID AAR85852 standard; peptide: 517 AA.

```

XX AC AAR85852;
XX 13-SEP-1996 (first entry)
XX DE WD-40 domain-contg. beta-TRCP protein.
XX DE WD40 repeat region; beta-transducin; protein-protein interaction; drug;
XX KM intracellular signalling; protein kinase C; homology; motif; modulator;
XX KM receptors of activated protein kinase; enzyme activity; isozyme; human.
XX OS Synthetic.
XX PN MO9521252-A2.
XX PD 10-AUG-1995.
XX PF 31-JAN-1995; 95WO-US01210.
XX PR 01-FEB-1994; 94US-0190802.
XX PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX PI Mochly-Rosen D, Ron D;
XX DR WPI; 1995-283772/37.
XX PT New WD-40 (beta-transducin)-derived polypeptide(s) - which alter the
XX PT activity of a protein, eg. protein kinase C, which interacts with a
XX PT protein contg. a WD-40 region.
XX PS Example 5; Page 80-82; 351pp; English.
XX XX
XX CC Proteins AAR85851-92 are protein which contain at least one WD-40 (also
XX CC called beta-transducin homologous) amino acid repeat motifs. The WD-40
XX CC regions are involved in protein-protein interactions between proteins
XX CC involved in intracellular signalling. An example of such an interaction
XX CC is between protein kinase C and receptors of activated protein kinase
XX CC (RACK), esp. RACK-1 (AAR85850). Proteins AAR85851-82 were isolated based
XX CC on homology with beta-transducin, whereas proteins AAR85882-92 were
XX CC isolated based on homology with the WD-40 consensus sequence (AAR85893).
XX CC The proteins were used to construct the peptides AAR84928-R85063 and
XX CC AAR85766-R85842. The peptides can be used to identify target proteins
XX CC contg. WD-40 motifs, as modulators of enzyme esp. isozyme, activity of
XX CC proteins involved in protein-protein interaction and to screen for drugs
XX CC that will affect protein-protein interaction involving WD-40 domains.
XX SO
Sequence 517 AA:

Query Match 85.1%; Score 2582.5; DB 16; Length 517;
Best Local Similarity 91.4%; Pred. No. 3,4e-243;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

OY 18 SIERECNNGEPPRKIIIEKNSLRQYNSCARLCINQETVCLASTAMTEMCVAKTKLAN 77
DB :||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13 aserecndrdepprkilieknlr-----tklan 42

OY 78 GTSSMIVRQRKLSAYEKELCYKFEOWSESDQVEFEHLISOMCHYOHGHTNSLYK 137
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
43 gtsmlvprkqlsanyskekelckvteqsecdqvefvehllismchghghnnylk 102

OY 136 PM/QDFTTALPARGLDHAENILSYDAKSLCAAEIVCKEMRYVTSDDGLMKRLIERMV 197
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
103 pmlqdfitalparglidhaenilslsydaksicsaelvckewyrvtsdgmllwkliermv 162

OY 198 RFDLSMRGLAERRGGOVLFFKNKPPDGNAPNSFRALYPRITODIETIESNMWCKGRSL 257
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 rtdslmrjlaerwggylfknhkppdgtktpnstyralypklidietiesnmwrcgrhsl 222

OY 258 QRHCRESKSGVYCLQYDDOKIYSGLDNNTIKIMDKNTLECKRKLTHGHTSVCLQYDE 317
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
223 qriheretskgyvclqyddqkivsgldntikimdkntleckrklthghgsvclqyde 282

```

```

OY 318 RVIIITGSSDSTFRVWDVNTGEMLNTLLIHCEAVLHLRFNNGMWTGSKDRSIAVMDASP 377
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 rviiltg-sdstfrvwdvntgemitllihceavlhrlfrnngmwtcskdrsiavmdasa 341

OY 378 TDILRLRVLGHRAAVVWVPFDDKYYIASGSDRIKXWNTSTGEFVTLNGHKGRIACLG 437
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
342 tdlrlrvlyghraavvwdfdkyyiasgsdrlkxwnstgefvtlngkhrgriacldq 401

OY 438 YRDLRVVSGSSDNTIRLMDIECGACLRVLGHEELVRCIRFDNKRRIYSGAYDGRIKWDL 497
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 yrdlrvvsgssdntirlmdiecgacrlvlgheelvrcirfdnkrrivsgaydgrikwld 461

OY 498 VAALDPRAPACTLCLRLVHSHGRVFRPLQDFEQLVSSSHDDPTLLIMDFLNDP 550
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
462 vaaldprapactlclrlvshshgrvfrplqdfelqvssshddptllimdfndp 514

RESULT 13
ID AAY96696 standard; Protein; 542 AA.
XX AC AAY96696;
XX XX
XX DT 26-SEP-2000 (first entry)
XX XX
XX DE Human E3 ubiquitin ligase.
XX XX
XX KM E3 ubiquitin ligase; beta-TRCP; F-box; WD protein; I-kappa-B; inhibitor;
XX KM nuclear factor kappa-B; NF-kappa-B; degradation; modulator; anti-viral;
XX KM anti-inflammatory; immunosuppressive; cytostatic.
XX OS Homo sapiens.
XX XX
XX PN W0200034447-A2.
XX PD 15-JUN-2000.
XX PF 10-DEC-1999; 99MO-US29371.
XX PR 10-DEC-1998; 98US-0210060.
XX PA (SIGN-) SIGNAL PHARM INC.
XX PA (YISS ) YISSUM RES & DEV CO.
XX PI Manning AM, Mercurio F, Amit S, Ben-eriah Y, Davis M, Hatzubai A;
XX PI Lavon I, Yaron A;
XX DR WPI; 2000-431294/37.
XX DR N-PSDB; AAA51228.
XX PT Polypeptide enhancing phosphorylated I-kappa-B ubiquitination useful for
XX PT treating disorder associated with NF-kappa-B activation e.g. cancer,
XX PT compising amino acid sequence of human E3 ubiquitin ligase or its
XX PT variant
XX PS Claim 1; Page 70-72; 77pp; English.
XX XX
XX CC This is human E3 ubiquitin ligase (E3), which is homologous to human
XX CC beta-TRCP, an F-box/WD protein family member. E3 enhances ubiquitination
XX CC of phosphorylated I-kappa-B, an inhibitor protein of nuclear factor
XX CC kappa-B (NF-kappa-B). Understanding I-kappa-B degradation via the
XX CC ubiquitin pathway is useful for identifying modulators of this process
XX CC for use in treating diseases associated with activation of NF-kappa-B. In
XX CC vitro analysis suggests that deletion of the F-box results in a protein
XX CC that functions as a dominant negative molecule in vivo. Transient
XX CC over-expression of delta-beta-TRCP (a deletion mutant) inhibited the
XX CC degradation of endogenous I-kappa-B-alpha in stimulated Jurkat cells,
XX CC resulting in accumulation of phosphorylated I-kappa-B-alpha. E3 can be
XX CC used to screen for modulators of NF-kappa-B activity. E3 and beta-TRCP
XX CC can be used to modulate NF-kappa-B to treat inflammatory diseases,
XX CC autoimmune diseases, cancer and viral infections.

```


Db 332 lmvtskdrslavwmasatdltrrvlyghraavnvddfdckylvasagdrllkwsts 391
 Qy 419 TCEFVRLNGHKRGIAQLQYRDRLVYSGSSDNTIRLMDIEGACIRVLEGEHELYRCIRF 478
 Db 392 tcefvrlngkhrgiaqlqyrdrlvsgssdntirlwdiecgacirvleghelvrclrf 451
 Qy 479 DNKRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCIRTVEHSGRVFRLODFEQIVSSSHD 538
 Db 452 dnkrivsgaydgkikvmdlvaaidprapastclrtlvhsgrvfrlqfdeqivssshd 511
 Qy 539 DTILWDFLNDPPAAQAEPSPSRRTYTYISR 569
 Db 512 dtllwdflnvppsaqnetrsprtytyisr 542
 RESULT 15
 AAM41994
 ID AAM41994 standard; Protein: 550 AA.
 XX AAM41994;
 AC AAM41994;
 XX 22-OCT-2001 (first entry)
 DT Human polypeptide SEQ ID NO 6925.
 DE Human polypeptide SEQ ID NO 6925.
 XX Human; nocotropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS Homo sapiens.
 PN WO200133312-A1.
 PD 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US34263.
 PF 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Huang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Dzmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB: AAI61150.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX Example 2; SEQ ID NO 6925; 10078pp; English.
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
 CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX Sequence 550 AA:
 SQ
 Query Match 78.6%; Score 2384.5; DB 22; Length 550;
 Best Local Similarity 79.0%; Pred. No. 8.1e-224;
 Matches 451; Conservative 48; Mismatches 41; Indels 31; Gaps 7;
 Qy 1 MDPRAEVLQKALKFNSSREDCNNGEPPRKILPEKNSLRQYNSCARLQNOEVCIA 60
 Db 9 mep-dsylviedkcltclms-----v-----slwlgcanlv---esmcsl 44
 Qy 61 S--TAMKTENCVAKTKLANGTSSMIVPKOKLSAYEKEKELCVKFEQMSSEDOVEYE 118
 Db 45 sclqsmprsvrcl--qngtssvsvsrkprsegnykckldcklyfdqwsedqyefve 101
 Qy 119 HLISQMGCHYGHINSLKLMLOPFTTALPAGLDIHANILSYLDAKSICAEIYCKE 178
 Db 102 hlismchghghlnslkmpldgrdftalpeghldhlaenllsyldarslcaaelvcke 161
 Qy 179 WYRTSDGMIMKRLIERMVRDTSIMRGLAERGMGYLFENKPPDGNAPPNSFYRALYRK 238
 Db 162 wgrvisegmllkklkiermvrtdpklwglseriywdqylfknrptdg--ppnsfyrlsyk 219
 Qy 239 ITQDIETIESNWRGRHSRLRHCRESKGVYCLQYDQKIVSGLRDNTIKWDKNTLE 298
 Db 220 itqdietsnwrgrhnlgrdqrsenskyvclgyddeklisqirdnsiklwdktsle 279
 Qy 299 CKRILTGHTGSVYCLQYDERVITITGSSDSTVRYWVDVNTGEMTLIHCEAVLHREFNNG 358
 Db 280 cikvlgtgsvclqydervivgssdstvrywvdvntgevtllhneavvlhlfng 339
 Qy 359 MMYVCSKDRSLAVWDMASPTDITLRLVYGHRAAVNVVDFDKRYIVSASGDRITKYWNIS 418
 Db 340 lmvtskdrslavwmasatdltrrvlyghraavnvddfdckylvasagdrllkwsts 399
 Qy 419 TCEFVRLNGHKRGIAQLQYRDRLVYSGSSDNTIRLMDIEGACIRVLEGEHELYRCIRF 478
 Db 400 tcefvrlngkhrgiaqlqyrdrlvsgssdntirlwdiecgacirvleghelvrclrf 459
 Qy 479 DNKRIVSGAYDGKIKVMDLVAAIDPRAPAGTLCIRTVEHSGRVFRLODFEQIVSSSHD 538
 Db 460 dnkrivsgaydgkikvmdlvaaidprapastclrtlvhsgrvfrlqfdeqivssshd 519
 Qy 539 DTILWDFLNDPPAAQAEPSPSRRTYTYISR 569
 Db 520 dtllwdflnvppsaqnetrsprtytyisr 550

Search completed: May 8, 2002, 10:50:24
 Job time: 111 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2002, 10:48:53 ; Search time 14.24 Seconds
(without alignments)
899.184 Million cell updates/sec

Title: US-09-601-168A-2

Perfect score: 3034
Sequence: 1 MDPAEAVLQKALKFMNSSE.....PAAQAPPPSPSTYYISR 569

Scoring table:
BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Dackfilest.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2582.5	85.1	517	1 US-08-190-802A-30	Sequence 30, Appl
2	2582.5	85.1	517	4 US-08-477-346-30	Sequence 30, Appl
3	520	17.1	587	3 US-08-899-578-2	Sequence 2, Appl
4	399	13.2	779	1 US-08-190-802A-32	Sequence 32, Appl
5	399	13.2	779	4 US-08-477-346-32	Sequence 32, Appl
6	354	11.7	409	2 US-08-283-917-3	Sequence 3, Appl
7	354	11.7	409	2 US-08-961-716-3	Sequence 3, Appl
8	354	11.7	409	2 US-08-283-917-9	Sequence 9, Appl
9	354	11.7	409	2 US-08-961-716-9	Sequence 9, Appl
10	339.5	11.2	409	1 US-08-190-802A-51	Sequence 51, Appl
11	339.5	11.2	409	4 US-08-477-346-51	Sequence 51, Appl
12	321.5	10.6	514	1 US-08-190-802A-66	Sequence 66, Appl
13	321.5	10.6	514	4 US-08-477-346-66	Sequence 66, Appl
14	318	10.5	422	1 US-08-190-802A-52	Sequence 52, Appl
15	318	10.5	422	4 US-08-477-346-52	Sequence 52, Appl
16	313.5	10.3	209	3 US-08-899-578-6	Sequence 6, Appl
17	306	10.1	704	1 US-08-188-582-5	Sequence 5, Appl
18	306	10.1	704	1 US-08-646-715-5	Sequence 5, Appl
19	305	10.1	704	1 US-08-190-802A-62	Sequence 62, Appl
20	305	10.1	704	2 US-08-308-818-3	Sequence 3, Appl
21	305	10.1	704	4 US-08-477-346-62	Sequence 62, Appl
22	299.5	9.9	1194	4 US-09-092-508-2	Sequence 2, Appl
23	298.5	9.8	1205	4 US-09-092-508-16	Sequence 16, Appl
24	298.5	9.8	704	1 US-08-188-582-18	Sequence 18, Appl
25	298.5	9.8	704	1 US-08-646-715-18	Sequence 18, Appl
26	285.5	9.4	798	1 US-08-190-802A-64	Sequence 64, Appl
27	285.5	9.4	798	1 US-08-190-802A-68	Sequence 68, Appl

28	285.5	9.4	798	2 US-08-308-818-2	Sequence 2, Appl
29	285.5	9.4	798	4 US-08-477-346-64	Sequence 64, Appl
30	285.5	9.4	798	4 US-08-477-346-68	Sequence 68, Appl
31	283.5	9.3	212	3 US-08-899-578-7	Sequence 7, Appl
32	283	9.3	713	1 US-08-190-802A-63	Sequence 63, Appl
33	283	9.3	713	4 US-08-477-346-63	Sequence 63, Appl
34	263	8.7	318	1 US-08-190-802A-33	Sequence 33, Appl
35	263	8.7	318	4 US-08-477-346-33	Sequence 33, Appl
36	260.5	8.6	375	4 US-09-063-743-1	Sequence 31, Appl
37	256	8.4	906	1 US-08-190-802A-31	Sequence 31, Appl
38	256	8.4	906	4 US-08-477-346-31	Sequence 31, Appl
39	246.5	8.1	343	4 US-09-063-743-5	Sequence 5, Appl
40	245.5	8.1	317	1 US-08-190-802A-27	Sequence 27, Appl
41	245.5	8.1	317	1 US-08-190-802A-41	Sequence 41, Appl
42	245.5	8.1	317	1 US-08-190-802A-47	Sequence 47, Appl
43	245.5	8.1	317	4 US-08-477-346-27	Sequence 27, Appl
44	245.5	8.1	317	4 US-08-477-346-41	Sequence 41, Appl
45	245.5	8.1	317	4 US-08-477-346-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1
US-08-190-802A-30
Sequence 30, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: MD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0860
TELEFAX: (415) 324-0960
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 517 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: BETA TRCP, Fig. 13
US-08-190-802A-30

Query Match 85.1%; Score 2582.5; DB 1; Length 517;
Best Local Similarity 91.4%; Pred. No. 5.5e-266;
Matches 487; Conservative 7; Mismatches 8; Indels 31; Gaps 2;

RESULT 2
 US-08-477-346-30
 Sequence 30, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WO-40 - Derived Peptides and Uses
 TITLE OF INVENTION: thereof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Morison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-0760
 TELEFAX: (202) 887-0763

Query Match	85.18;	Score 2582.5;	DB 4;	Length 517;
Best Local Similarity	91.48;	Pred. No. 5.5e-266;		
Matches 487; Conservative	7;	Mismatches 8;	Indels 31;	Gaps 21

QY	18	SSSEEDCNNGEPPRKITPEKNSLRBTYSNARCJLNGEYVCLASTMKRENOVAKTKIAN	77
Db	13	ASREEDCNDEPPRKITPEKNTLRO-----TKIAN	42
QY	78	GTSSMIVPKOKLSASYEKEKLCVKYFEQWSESDQVEFVEHLISQMCYHGHINSYLK	137
Db	43	GTSSMIVPKOKLSANEYKEKLCVKYFEQWSESDQVEFVEHLISPMCHYGHINITYLK	102
QY	138	PMIORPITALPARGIDHIAENILSYLDASLCAELUYCKEWMYRTSDGMKKKILEMY	197
Db	103	PMIORPITALPARGIDHIAENILSYLDASLCSAEUYCKEWMYRTSDGMKKKILEMY	162
QY	198	RTDSLWRGLAERBGQVLYFNKPPDDCNAPNSFYALYPKIIIDDIETESNMGRSHL	257
Db	163	RTDSLWRGLAERBGQVLYFNKPPDDKTPNPSFYALYPKIIIDDIETESNMGRSHL	222
QY	258	QRHCSESESKVYCYLQYDDQKIVYSGLRDNTIKIMOKNTLECKRLLTGHTSVYLCLOYDE	317
Db	223	QRHCSESESKVYCYLQYDDQKIVYSGLRDNTIKIMOKNTLECKRLLMGHTSVYLCLOYDE	282
QY	318	RVITIGSSDSTYRVWMDVNTGEMLNTLLHHC EAVLHLRFNNGMAYTCSKDRSIAVWDASP	377
Db	283	RVITIG -SDSTYRVWMDVNTGEMLNTLLHHC EAVLHLRFNNGMAYTCSKDRSIAVWDASA	341
QY	378	TDITLRYLYGHRAAVNVNPDQKIVYSASGDTIKYMWYSTCEYVRLNGHKGINCLQ	437
Db	342	TDITLRYLYGHRAAVNVNPDQKIVYSASGDTIKYMWYSTCEYVRLNGHKGINCLQ	401
QY	438	YRDLRVSSGSNDTIRLMDIECGALCYLVEGHEBELVRCIRFENKRIYSAGVAGDKITAWDL	497
Db	402	YRDLRVSSGSNDTIRLMDIECGALCYLVEGHEBELVRCIRFENKRIYSAGVAGDKITAWDL	461
QY	498	VVALDPRADGICLIRLYEHSNGVFPYLDQDEQIYSSSHDTPILIMPLNDP	550
Db	462	VVALDPRADGICLIRLYEHSNGVFPYLDQDEQIYSSSHDTPILIMPLNDP	514

RESULT 3
 US-08-899-578-2
 Sequence 2, Application US/08899578
 Patent No. 6087153
 GENERAL INFORMATION:
 APPLICANT: Greenwald, Iva
 APPLICANT: Hubbard, E. Jane
 TITLE OF INVENTION: SEL-10 AND USES THEREOF
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESS: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/899,578
 FILING DATE: 24-JUL-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 0575/53200/JFW/ACC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 278-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 587 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-899-578-2

Query Match 17.1%; Score 520; DB 3; Length 587;
 Best Local Similarity 28.8%; Pred. No. 2,3e-46;
 Matches 150; Conservative 78; Mismatches 222; Indels 70; Gaps 17;

72 KTKANGSTSMI-----VPRKOR--LSASYEKEL-----CVKFEQWSESDOYE 115
 35 ESSYSSSSSSSYNAADKISSRPLQHKILDLASPSRNNDLNPRVHLLALFKDLSSADMD 94
 116 FVEHLISQMCYHGHINSYLRKMLQDPITLAPRGDLHIAENILSYDLAKSLCAAEIV 175
 95 AFPRLLQESNMNTIRQLRAIIEPHQDFLCLPV---ELGKMLHLNLGYDILLVYAOV 150
 176 CKERYRTSDGMLMKL-IERM-----RTDSLMRGLAERRGQGYLFKNRPPDGNAP 227
 151 SKNNKLISEIDKIMKSLGVEEFKHHPDTRVGTAGMOGTIAAG-----VTPDHITP 203
 228 PN-SFYALYPKIIDI-----ETIESNRCGRHSIORHCRSETSKGYCL 273
 204 CDLAVHFKLQKRGDIFERRADKSRRLADKIEKNANPIMGSAV-LGHEDHIVTMC 262
 274 QYDQKIVSGLRDNTIKIMDKNTECKRILTGHTGSLVLCIODE--RVITGSSDSIVRV 331
 263 QIHDDVATVGTSDMTLKWICIDKGEVMTYVGHGTGVTWSIOCGRYIVSGSTDRIVKV 322
 332 MDVNTGEMNLNLIHCEAVLHLRFNNGMVTGSKDRSIAVDMASPTDITLRVYVGHRA 391
 333 WSYVDGSLHTLQGHSTVRCMAAGSLVYGSDDTLRLWVDVSGRHLA---TLGHHA 379
 392 AVNVVDEDDKIYVASGDRITIKVNTSTCEVFTRLNGHKGIGIACLYRDR--LVVSGSSD 449
 380 AVRCVQFGTIVVSGYFVYKINNAHTGRCIRLTGTNNHNVSLPESERSIYCSGSLD 439
 450 NTRILMDI---ECGACLRVLEGHHELVRCIFDKKRIYSGVADKIKVMDLVALDRAP 506
 440 TSIKRWDTREGECEVALLQGHSTLSTGMLRGNIILVSCNADSHVRWDI-----H 491
 507 AGTLCLRTLVESGKRVFLQ--FDEFQIVSSSHDDITLIMD 545
 492 EGT-CVHMLSGHRSRATISLQWFGGRMVAITSSDDGIVKIMD 530

RESULT 4
 US-08-190-802A-32
 Sequence 32, Application US/08190802A
 Patent No. 5519003
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Theoreof
 NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Dehlinger & Associates
 STREET: P.O. Box 60850
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94306-0850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/190,802A
 FILING DATE: 01-FEB-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Fadian, Gary R.
 REGISTRATION NUMBER: 33,875
 REFERENCE/DOCKET NUMBER: 8600-0139
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-0880
 TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 32:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 779 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: CD4 / CD20 protein, Fig. 15
 US-08-190-802A-32

Query Match 13.2%; Score 399; DB 1; Length 779;
 Best Local Similarity 24.3%; Pred. No. 2.8e-33;
 Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

56 TVCLASTAMKTCNVAKTKLANGSTSMIVPKQKLSASYEKEL-----ELCVYFEQW 108
 194 TPLATTTKTINN-----NNNIADLIESKDISISPEYLSDEIFSAJNNNPHAYFK-- 244
 109 SESDOVEFEVHLSQMCYHGHINSYLRKMLQDPITLAPRGDLHIAENILSYDLAKS 168
 245 -----NLFRIVANNDRELSDLGTLIKDNLRDLITSLP---ELSLKFTNLOPED 294
 169 LCAAEIVCKEWR-VTSDGMLMKLI-ERMVRTDSLMRGLAERRGQGYLFKNRPPDGN 225
 295 IINSLVGSQWNNKIIKRSISLMKLLISENFV-----SPKGF 331
 226 APPNSFYALYPKIIDI-----IET--TESNRCGRHSIORHCRSETSKGYCLOYD 276
 332 NSLNLKLSQKRYPKLSOODRLSLFLENIFLKNWYMPKVPQDTTLRGHMTSVITCLOPE 391
 277 DQKIVSGLRDNTIKIMDKNTECKRILTGHTGSLVLCIODE--RVITGSSDSIVRVYDN 335
 392 DNVYITGADKMTIRVDSINKKFLDLSGHDGVMALKAHGSILVSGSTDRIVRWDIK 451
 336 TGEMLNTLIHCEAVLHLRFNNGMVTGSKDRSIAVDMASPTDITLRVYVGHRAAVV 395
 452 KG-----CCT-----HVFEGHNSVRC 468
 396 VDEPD---KIYVASGDRITIKVNT-----STCE---FVFTLNG 428
 469 LDIVEYKNIKIYVYTGSDNLTILHWKLPKESSVPDHEEDHYPVFTPEBNPYFVGLNG 528
 429 HKRGIACTORYRDLVWVGSSDNTIRLMDIEGACLRVLEGHHELVRCIRFDN--KRIYSG 486
 529 HMASVRIYSGHNTIVSGSDNTILWVDMQMCIVILSGHTBRIVSTIYDHRKRCISA 588
 487 AYDKIKRVMDL-----VAALDPRAPAGTL--CLRTLVESGKRVFLQDFQIVSS 535

Db 589 SMDTIRLWDLENINWNGECASYATNSASPCAKILGAMTYLQGHIALVGLLSDKFLVSA 648
QY 536 SHDDTILIMDFLNDPAQAQEPSPRSRTYTY 566
Db 649 AADGSIRGMD-AND-----YSRKFSY 668

RESULT 5
US-08-477-346-32

; Sequence 32, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: The use of
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477, 346
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487, 072
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: CDC4 / CDC30 protein, Fig. 15
; US-08-477-346-32

Query Match 13.2%; Score 399; DB 4; Length 779;
Best local similarity 24.3%; Pred. No. 2.8e-33;

Matches 139; Conservative 80; Mismatches 196; Indels 156; Gaps 20;

QY 56 TVCIASHTAKTENCQVAKTKIANGTSSMIVPKOKLSASYEKER-----ELCVYFEQW 108
Db 194 TPLAKTKTINN-----NNNIADLISKSDSIISPELSDIISATNNNIPHAFFK-- 244
QY 109 SESDQVEFEHLISOMCHQHGHSYAKPMQORDITLAPRGDHTAENILSYLDAS 168
Db 245 -----NLLFRLVANDRSELSDLGLIDNLRDLITSLP-----ELSLKTFNYLQFED 294
QY 169 LCAAEIVCKEWR-VTSDMLMKLI--ERAVRTDSLWGLAERGWQGYLFKNKPPDGN 225
Db 295 IINSLGVSQNMKKIIRKISLWKKLISINFY-----SPKGF 331

QY 226 APPNSFYRLYPEIID-----IET--IESNRCGRHSIORINCRSETSGVYCLQYD 276
Db 332 NSLNLKLSQRYPKLSOODRLRLSFLLENIFLTKWNPYKFPQRTLLGHMTSVICLOPE 391
QY 277 DOKIVSGLRNRTIKIWDKNTLECKRLITLGHSGVLCLOQDE-RVIITGSSDSIVRWYDN 335
Db 392 DNYVITGADDMKIRYDTSINKKFLLOLSGHDGVALAKYAHGGLIVSGSTDRTPVWMIK 451
QY 336 TGEMLNTLIHCEAVLHLRFNNGMWTCSKDRSIAMVDMASPTDILRLRVLIGHRAAVN 395
Db 452 KG-----CCT-----HVFEGHSTYRC 468
QY 396 VDFDD-----KIYVASGDTIKYWN-----STCE-----FVPTLNG 428
Db 469 LDIVKRNKIKYVTSGRDNTLHWKLPKSSVPDGHGEHDYPLVFTPEENPYFVGLRG 528
QY 429 HKRGIACLOIRLIVYSGSSDNTIRLMDIEGACLVLEGHELERCIRFEDN-KRIYS 486
Db 529 HMASVTVSGHGNIVYSGSDNTLIVMDYAKMKCLYILSGHTRILYITIDHERRKICISA 588
QY 487 AIDGKIKWDL-----VALDPRAPAGTL-CLRTLVHSGVRPLQFDEFOIVSS 535
Db 589 SMDTIRLWDLENINWNGECASYATNSASPCAKILGAMTYLQGHIALVGLLSDKFLVSA 648
QY 536 SHDDTILIMDFLNDPAQAQEPSPRSRTYTY 566
Db 649 AADGSIRGMD-AND-----YSRKFSY 668

RESULT 6

; Sequence 3, Application US/08283917
; Patent No. 5849557

; GENERAL INFORMATION:
; APPLICANT: ADACHI, HIDEKI
; APPLICANT: TSUJIMOTO, MASAFUMI
; APPLICANT: INOUE, KEIZO
; APPLICANT: ARAI, HIROYUKI
; TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
; TITLE OF INVENTION: AND GENE THEREOF
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER &
; ADDRESS: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,917
; FILING DATE: 03-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 209943/1993
; FILING DATE: 03-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5849557man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2292-030-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TEXT: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 409 amino acids
; TYPE: amino acid

```

; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
US-08-283-917-3

```

Query Match	11.7%;	Score 354;	DB 2;	Length 409;
Best Local Similarity	29.2%;	Pred. No. 5.8e-29;		
Matches	87;	Conservative	57;	Mismatches 110; Indels 44; Gaps 8

```

QY 280 IYSGLEDTNIIKIMDNTECKRIILTHGYSLOCLOYDE--RYIINGSSISPTRYAWDNTG 337
Db 122 MWSASEDATIKWMDIETGDEFERTKGTBDYODISDPHSGKILLASCSADMIKIMLDPGFG 187
QY 338 EMILTLIIHCSEAVLHFR--NNGMAMVTCSDRSTAYWDMASPDITLRRVLVGHRAAVNY 395
Db 182 ECIRTHMGHDHNVSSVALPMPGDHIVASRSRDKTIKMEVOTGYCV---KTFGHEWVRM 238
QY 396 V--DPPDCKIVASGGRITKJMNSTCEPFRYLGNHRCJACLOYDR-----441
Db 239 VRPNODSTLASCNDQYRIVWVAATKECKEALREHNHVEECISMAPESSYSISEATGS 296
QY 442 -----LVVSGSSDNTIRLMDIECGACRLVLESGHELVNCFRFDN--KRIYSGAYD 488
Db 299 ETYKSGKPGFELLGSRDRTIKMMDVSTGCMILMTLVGHDMMWVAGVLFHSGKFFLISCADD 356
QY 490 GKIKWDLVALDPRPAGTCLIRLYHSGCRVRLQDFDEF--QIYSSHDITILIMD 545
Db 359 KTLKMWY-----KKKRCMKTLNHNHFVTSLDHFHTATAYVTGSGVDOTYAVME 407

```

RESULT 7
 US-08-961-716-3
 ; Sequence 3, Application US/08961716
 ; Patent No. 5880272
 ; GENERAL INFORMATION:
 APPLICANT: ADACHI, HIDEKI
 APPLICANT: TSUJIMOTO, MASAFUMI
 APPLICANT: INOUE, KEIJO
 APPLICANT: ARAI, HIROYUKI
 TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
 TITLE OF INVENTION: AND GENE THEREOF
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: OELON, SEIVAK, MCCLELLAND, MATER &
 ADDRESSEE: NESTLADT, P.C.
 STREET: 1755 S. Jefferson Davis Highway, Suite 400
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,716
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/283,917
 FILING DATE: 03-AUG-1994
 APPLICATION NUMBER: JP 209943/1993
 FILING DATE: 03-AUG-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Obilon, No. 5880272man F.
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2292-030-0
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220
 TELE: 248855 OPAT UR
 INFORMATION FOR SEQ. ID NO:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 409 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus

Query Match	11.78;	Score 354;	DB 2;	Length 409;
Best Local Similarity	29.28;	Pred. No. 5.8e-29;		
Matches 87; Conservative	57;	Mismatches 110;	Indels 44;	Gaps 8

```
OY      280   IVGLRNTIKIMDKNLTCECRILITGHTAGSLCLOYDE--RVIIIGSSDSIVRWAVDNTG    369
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      122   MVASSEDATIKWMDVETQDFEFLTKGHDTVDIUSIFSGKLLACSCADMTIKLMDFOGE    181
OY      338   EMLTTLIHHCVALHLRF--NNGMAYTCSDRSIAVMWDASPFDDITLRVLGVRAALVN    395
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      182   ECIRTMIGHDNHNVSAIWPNDNHIVASSBCKTIKKMEVOTGCV---KTTPGHEWRM     238
OY      396   V--DEFODKYVASGGRTIKVWNSTCFEFVTLLNGHKRGICAOXYDR-----            441
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      239   VRPNOSTILLASCNSNOQIVRWVVATIKCCAKELNEHENHYVECSISMAPESSYSISEATGS    298
OY      442   -----LVYSSSDNTFRAMDIECGACLRLVBSEHELVACIFDN--KRIVSAYD    489
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      299   ETKKSGPGPFFLLSGSRDXTIKAMPVSFGOMLMTLVGDHMWVRGVLFGHSKGKFLLISCADD    358
OY      490   GKIKWIDLVAAIDRAPPACTLLCTRPLFVGSGRPVRILODFEE--OIYSSSHDGTILIIM    545
          :| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      359   KTLRWMIY-----KKRKCMKLNAHNHFYSLDPHKTATAYVTTSVDQTAWME    407
```

RESULT 8
US-08-283-917-9
: Sequence 9, Application US/08283917
: Patent No. 5849557
: GENERAL INFORMATION:
: APPLICANT: ADACHI, HIDEKI
: APPLICANT: TSUJIMOTO, MASAFUMI
: APPLICANT: INOUE, KEIZO
: APPLICANT: ARAI, HIROYUKI
: TITLE OF INVENTION: OXIDIZED PHOPHOLIPID DEGRADING ENZYME
: TITLE OF INVENTION: AND GENE THEREOF
: NUMBER OF SEQUENCES: 31
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
: ADDRESSEE: NEUSTADT, P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/283,917
: FILING DATE: 03-AUG-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 209943/1993
: FILING DATE: 03-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NO. 5849557/man F.


```

1      RESULT 10
2      US-08-190-802A-51
3      ; Sequence 51, Application US/08190802A
4      ; Patent No. 5519003
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Mochly-Rosen, Darla
7      ; APPLICANT: Ron, Dorit
8      ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
9      ; TITLE OF INVENTION: Thereof
10     ; NUMBER OF SEQUENCES: 265
11     ; CORRESPONDENCE ADDRESS:
12     ; ADDRESSEE: Delhinger & Associates
13     ; STREET: P.O. Box 60850
14     ; CITY: Palo Alto
15     ; STATE: CA
16     ; COUNTRY: USA
17     ; ZIP: 94306-0850
18     ; COMPUTER READABLE FORM:
19     ; MEDIUM TYPE: Floppy disk
20     ; COMPUTER: IBM PC compatible
21     ; OPERATING SYSTEM: PC-DOS/MS-DOS
22     ; SOFTWARE: PatentIn Release #1.0, Version #1.25
23     ; CURRENT APPLICATION DATA:
24     ; APPLICATION NUMBER: US/08/190,802A
25     ; FILING DATE: 01-FEB-1994
26     ; CLASSIFICATION: 530
27     ; ATTORNEY/AGENT INFORMATION:
28     ; NAME: Fabian, Gary R.
29     ; REGISTRATION NUMBER: 33,875
30     ; REFERENCE/DOCKET NUMBER: 8600-0139

```

RESULT 11
 US-08-477-346-51
 ; Sequence 51, Application US/08477346
 ; Patent No. 6262023
 ;
 GENERAL INFORMATION:
 ;
 APPLICANT: Mochly-Rosen, Daria
 ;
 APPLICANT: Ron, Dorit
 ;
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ;
 TITLE OF INVENTION: Theoreof
 ;
 NUMBER OF SEQUENCES: 265
 ;
 CORRESPONDENCE ADDRESS:
 ;
 ADDRESSEE: Morrison & Foerster
 ;
 STREET: 2000 Pennsylvania Avenue, NW
 ;
 CITY: Washington
 ;
 STATE: DC
 ;
 COUNTRY: USA
 ;
 ZIP: 20006-1812
 ;
 COMPUTER READABLE FORM:
 ;
 MEDIUM TYPE: Floppy disk
 ;
 COMPUTER: IBM PC compatible
 ;
 OPERATING SYSTEM: PC-DOS/MS-DOS
 ;
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 ;
 CURRENT APPLICATION DATA:
 ;
 APPLICATION NUMBER: US/08/477, 346
 ;
 FILING DATE: 07-JUN-1995
 ;
 CLASSIFICATION: 514
 ;
 PRIOR APPLICATION DATA:
 ;
 APPLICATION NUMBER: 08/487, 072
 ;
 FILING DATE: 07-JUN-1995
 ;
 ATTORNEY/AGENT INFORMATION:
 ;
 NAME: MORASHIGE, KATE H.
 ;
 REGISTRATION NUMBER: 29,959
 ;

RESULT 12
 US-08-190-802A-66
 ; Sequence 66, Application US/08190802A
 ; Patent No. 5519003
 ; GENERAL INFORMATION:
 ; APPLICANT: Mochly-Rosen, Daria
 ; APPLICANT: Ron, Dorit
 ; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 ; TITLE OF INVENTION: Theoreof
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Delhinger & Associates
 ; STREET: P.O. Box 60850
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94306-0850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/190,802A
 ; FILING DATE: 01-FEB-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fadian, Gary R.
 ; REGISTRATION NUMBER: 33,875
 ; REFERENCE/DOCKET NUMBER: 8600-0139
 ; TELECOMMUNICATION INFORMATION:
 ;

RESULT 14
US-08-190-802A-52
Sequence 52, Application US/08190802A
Patent No. 5513003

GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WO-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dellinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-0850

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,802A
FILING DATE: 01-FEB-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fabian, Gary R.
REGISTRATION NUMBER: 33,875
REFERENCE/DOCKET NUMBER: 8600-0139
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960
 INFORMATION FOR SEQ ID NO: 52:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 422 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: MD6, Fig. 35
 US-08-190-802A-52

Query Match 10.5%; Score 318; DB 1; Length 422;
 Best Local Similarity 24.4%; Pred. No. 4.1e-25;
 Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

QY 103 KYFEQMSSEDOVEF-----VEHLISQMGHYGHINSYLMKPMLOPDTALPAR 151
 DB 4 KDFETWLDNISVTFSLMDLQKNETLDHLISLGAVALRHLNNLETLKRDPLKLP- 62
 QY 152 GLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKLIERVATDSLMGLAERG 211
 DB 63 ---ELSFYLLKMDPQTLTCLCLVSKQKRVIS-----ACTEVMQACKNLG 106
 QY 212 WQOYLKKNPPGNAAPPNSFYRALYPIIODETIESNMRCGRHSIORHCRSETSKGY 271
 DB 107 W-----QIDDSVQDSILHMKRYLYKAILRMKOLED-----HEAFETSS--- 143
 QY 272 CLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVITGSSDSIVRV 331
 DB 144 -----LIGHSARYALYYKDGCLCTGSDLSAKL 172
 QY 332 MDVNTGEMLNTL-IHCEAVLHLRFNNGMMVTCSKRSIAVMDASPTDITLRLVYGR 390
 DB 173 WDVSTGQCVYGIQHTCAAV---KFDEQKLVTSFQNTVAQWEMSSGARTQHR---GHT 226
 QY 391 AAVNVDPEDK--YIVSASGDRITKVMNTSTCEPVRTLNGHKRG-----ACLO 437
 DB 227 GAVFSVYSDLDLIVSGADFAVKWALSAGTCLNLTGHTTEWTKVYLQCKVKSLLH 286
 QY 438 YRDLVYSGSSDNTIRLW---DIEGACLRVLEGHEELVRCIR---FNNKRIYS--- 485
 DB 287 SPGDYILSADKYELKIMPIRGREINC-KCLKTLYSVSDRSICLOPRLHFDGKIYVSSAL 345
 QY 486 GAYDGKIKVMDLVALDPRAPAGTLCRLTVHSGRVRFLOFD 528
 DB 346 GLYQWDFASYDILRVIKTPEVANLALL-----GFGDVFALLFD 383

RESULT 15
 US-08-477-346-52
 Sequence 52, Application US/08477346
 Patent No. 6262023
 GENERAL INFORMATION:
 APPLICANT: Mochly-Rosen, Daria
 APPLICANT: Ron, Dorit
 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
 TITLE OF INVENTION: Theoreof
 NUMBER OF SEQUENCES: 265
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Morrison & Foerster
 STREET: 2000 Pennsylvania Avenue, NW
 CITY: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/477,346
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/487,072
 FILING DATE: 07-JUN-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 2550-0025.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 INFORMATION FOR SEQ ID NO: 52:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 422 amino acids
 TYPE: amino acid
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 INDIVIDUAL ISOLATE: MD6, Fig. 35
 US-08-477-346-52

Query Match 10.5%; Score 318; DB 4; Length 422;
 Best Local Similarity 24.4%; Pred. No. 4.1e-25;
 Matches 113; Conservative 69; Mismatches 161; Indels 120; Gaps 16;

QY 103 KYFEQMSSEDOVEF-----VEHLISQMGHYGHINSYLMKPMLOPDTALPAR 151
 DB 4 KDFETWLDNISVTFSLMDLQKNETLDHLISLGAVALRHLNNLETLKRDPLKLP- 62
 QY 152 GLDHAENILSYLDAKSLCAELVCKEYRYTSDGMLMKLIERVATDSLMGLAERG 211
 DB 63 ---ELSFYLLKMDPQTLTCLCLVSKQKRVIS-----ACTEVMQACKNLG 106
 QY 212 WQOYLKKNPPGNAAPPNSFYRALYPIIODETIESNMRCGRHSIORHCRSETSKGY 271
 DB 107 W-----QIDDSVQDSILHMKRYLYKAILRMKOLED-----HEAFETSS--- 143
 QY 272 CLQYDDQKIVSGLRDNTIKIMDKNTLECKRILTGHTGVSVCLOYDERVITGSSDSIVRV 331
 DB 144 -----LIGHSARYALYYKDGCLCTGSDLSAKL 172
 QY 332 MDVNTGEMLNTL-IHCEAVLHLRFNNGMMVTCSKRSIAVMDASPTDITLRLVYGR 390
 DB 173 WDVSTGQCVYGIQHTCAAV---KFDEQKLVTSFQNTVAQWEMSSGARTQHR---GHT 226
 QY 391 AAVNVDPEDK--YIVSASGDRITKVMNTSTCEPVRTLNGHKRG-----ACLO 437
 DB 227 GAVFSVYSDLDLIVSGADFAVKWALSAGTCLNLTGHTTEWTKVYLQCKVKSLLH 286
 QY 438 YRDLVYSGSSDNTIRLW---DIEGACLRVLEGHEELVRCIR---FNNKRIYS--- 485
 DB 287 SPGDYILSADKYELKIMPIRGREINC-KCLKTLYSVSDRSICLOPRLHFDGKIYVSSAL 345
 QY 486 GAYDGKIKVMDLVALDPRAPAGTLCRLTVHSGRVRFLOFD 528
 DB 346 GLYQWDFASYDILRVIKTPEVANLALL-----GFGDVFALLFD 383

Search completed: May 8, 2002, 10:50:44
 Job time: 111 sec

